

A NOVEL FACIAL EXPRESSION RECOGNITION METHOD USING BI-DIMENSIONAL EMD BASED EDGE DETECTION

By

Zijing Qin

A Thesis

Submitted to the

Faculty of the Graduate School

of

Western Carolina University

in Partial Fulfillment of

the Requirements for the Degree

of

Master of Science in Technology

Committee:

\_\_\_\_\_ Director

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_ Dean of the Graduate School

Date: \_\_\_\_\_

Spring 2010  
Western Carolina University  
Cullowhee, North Carolina

A NOVEL FACIAL EXPRESSION RECOGNITION METHOD USING BI-DIMENSIONAL EMD BASED EDGE DETECTION

A thesis presented to the faculty of the Graduate School of Western Carolina University in partial fulfillment of the requirements for the degree of Master of Science in Technology.

By

Zijing Qin

Director: James Z. Zhang, Ph.D.  
Associate Professor  
Department of Engineering and Technology

April 2010

©2010 by Zijing Qin

This thesis is dedicated to my parents who have supported me all the way since the beginning of my studies.

This thesis is also dedicated to my grandfather who passed way three years ago. He has always been a great source of motivation and inspiration.

## ACKNOWLEDGEMENTS

I am greatly thankful to my major advisor, Dr. James Z. Zhang, whose guidance, patient encouragement, and support from the initial to the final level enabled me to develop an understanding of the subject and the writing of this thesis. I am also grateful for Dr. Robert D. Adams and Dr. Peter C. Tay's insightful criticisms and advice. Last but not least, I would like to thank Dr. Linda A. Camras from DePaul University who generously provided us reference facial expression images. Without any of those people's help, this thesis would not have been possible.

## TABLE OF CONTENTS

Acknowledgements . . . . .	iv
List of Tables . . . . .	vii
List of Figures . . . . .	viii
Abstract . . . . .	x
CHAPTER 1. Introduction . . . . .	4
1.1 Facial Expression Detection . . . . .	4
1.2 Edge Maps For Facial Expression Detection . . . . .	5
1.3 Contributions . . . . .	6
1.4 Structure of the Thesis . . . . .	6
CHAPTER 2. Literature Review . . . . .	7
2.1 Facial Expression Recognition . . . . .	7
2.1.1 The Facial Action Coding System . . . . .	7
2.1.2 Facial Expression Measurement . . . . .	9
2.1.3 Impact of Cultural Differences On Facial Expression . . . . .	10
2.2 Edge Detection . . . . .	11
2.2.1 Edge Detection Overview . . . . .	11
2.2.2 Predominant Edge Detection Methods . . . . .	13
2.3 Empirical Mode Decomposition Method . . . . .	17
CHAPTER 3. Methodology . . . . .	20
3.1 Overview of Method . . . . .	20
3.2 Facial Edge Map Extraction . . . . .	20
3.2.1 BEMD Method . . . . .	21
3.2.2 Mathematical Morphology . . . . .	22
3.3 Facial Expression Recognition Using Edge Map . . . . .	24
3.3.1 Facial Expression Recognition Criteria . . . . .	25
3.3.2 Determination of Facial Expression Recognition Criteria . . . . .	26
3.3.3 Facial Expression Recognition Procedures . . . . .	30
3.4 Facial Expression Blind Detection . . . . .	32

CHAPTER 4. Results . . . . .	34
4.1 The Japanese Female Facial Expression (JAFFE) Database . . . . .	34
4.2 Taiwanese Facial Expression Image Database (TFEID) . . . . .	39
4.3 Culture Differences and Facial Expressions . . . . .	41
4.4 Blind Detection For Both of The Databases . . . . .	43
4.5 Summary of Results . . . . .	45
CHAPTER 5. Conclusions and Future Work . . . . .	46
5.1 Restatement of the Problem and Objectives . . . . .	46
5.2 Conclusions . . . . .	46
5.3 Future Work . . . . .	47
Bibliography . . . . .	48
Appendix . . . . .	
APPENDIX A. Collected Data . . . . .	55

## LIST OF TABLES

2.1	Sample subset of AUs . . . . .	8
3.1	Four facial features for facial recognition . . . . .	25
3.2	Criteria for expression recognition . . . . .	30
4.1	Categorical success rate for JAFFE . . . . .	35
4.2	Percentage changes of D2 and D4 in JAFFE . . . . .	37
4.3	Detection rate for 6 facial expressions of JAFFE . . . . .	39
4.4	Categorical success rate for TFEID . . . . .	39
4.5	Detection rate for 6 facial expressions of TFEID . . . . .	40
4.6	D1-D4 values of one of the Japanese model . . . . .	42
4.7	Categorical success rate of the blind-detection algorithm tested on the JAFFE . . . . .	44
4.8	Detection rate for 6 facial expressions of the blind-detection algorithm tested on the JAFFE . . . . .	44
4.9	Categorical success rate of the blind-detection algorithm tested on the TFEID . . . . .	44
4.10	Detection rate for 6 facial expressions of the blind-detection algorithm tested on the TFEID . . . . .	45
A.1	Raw data of facial expression metrics tested on JAFFE . . . . .	55
A.2	Raw data of facial expression metrics tested on TFEID . . . . .	60
A.3	Raw data of facial expression metrics blind detection tested on JAFFE . . . . .	68
A.4	Raw data of facial expression metrics blind detection tested on TFEID . . . . .	74

## LIST OF FIGURES

2.1	Recognition example . . . . .	10
3.1	Overall process . . . . .	21
3.2	Examples using basic rectangular shape . . . . .	22
3.3	BEMD of Lena image: (1) Original image, (2)1 <sup>st</sup> IMF, (3)2 <sup>nd</sup> IMF, (4)3 <sup>rd</sup> IMF, (5)4 <sup>th</sup> IMF, (6)5 <sup>th</sup> IMF, (7)6 <sup>th</sup> IMF, (8)Residue . . . . .	23
3.4	(1) The 1 <sup>st</sup> BIMF of Lena image (2) Converting the grey scale image to binary image (3) Edge results after applying mathematical morphology . . . . .	24
3.5	Edge images obtained with different critical values in morphology operation: (1) 20 pixels (2) 60 pixels (3) 40 pixels . . . . .	24
3.6	Facial skeleton and characteristic distances . . . . .	25
3.7	Neutral facial expression: (1) Reconstructed image from BIMFs, (2) The 1 <sup>st</sup> BIMF, (3) Extracted edge image with bounding boxes . . . . .	26
3.8	”Happy” and edge map for recognition . . . . .	27
3.9	”Sad” and edge map for recognition . . . . .	27
3.10	”Fear” and edge map for recognition . . . . .	28
3.11	”Angry” and edge map for recognition . . . . .	28
3.12	”Disgust” and edge map for recognition . . . . .	29
3.13	”Surprise” and edge map for recognition . . . . .	29
3.14	Comparison with predominant edge detection methods . . . . .	31
3.15	Procedures and thresholds for detection . . . . .	32
3.16	Strategy examples for blind detection for mouth area . . . . .	33
3.17	Strategy examples for blind detection for mouth area . . . . .	33
4.1	Categorical detection results for JAFFE . . . . .	35
4.2	Detection results for ”Happy”, ”Sad”, ”Fear” . . . . .	36
4.3	Detection results for ”Angry” of JAFFE . . . . .	37
4.4	Detection results for ”Surprise” and ”Disgust” of JAFFE . . . . .	38
4.5	Categorical detection results for TFEID . . . . .	40

4.6	Detection results for "Happy", "Sad", and "Fear" of TFEID . . . . .	41
4.7	Detection results for "Angry" of TFEID . . . . .	42
4.8	Detection results for "Disgust" and "Surprise" of TFEID . . . . .	43
4.9	Japanese Female Facial Expressions . . . . .	43

## ABSTRACT

### A NOVEL FACIAL EXPRESSION RECOGNITION METHOD USING BI-DIMENSIONAL EMD BASED EDGE DETECTION

Zijing Qin, M.S.T.

Western Carolina University (April 2010)

Director: James Z. Zhang, Ph.D.

Facial expressions provide an important channel of nonverbal communication. Facial recognition techniques detect people's emotions using their facial expressions and have found applications in technical fields such as Human-Computer-Interaction (HCI) and security monitoring. Technical applications generally require fast processing and decision making. Therefore, it is imperative to develop innovative recognition methods that can detect facial expressions effectively and efficiently.

Traditionally, human facial expressions are recognized using standard images. Existing methods of recognition require subjective expertise and high computational costs. This thesis proposes a novel method for facial expression recognition using image edge detection based on Bi-dimensional Empirical Mode Decomposition (BEMD). In this research, a BEMD based edge detection algorithm was developed, a facial expression measurement metric was created, and an intensive database testing was conducted. The success rates of recognition suggest that the proposed method could be a potential alternative

to traditional methods for human facial expression recognition with substantially lower computational costs. Furthermore, a possible blind-detection technique was proposed as a result of this research. Initial detection results suggest great potential of the proposed method for blind-detection that may lead to even more efficient techniques for facial expression recognition.

## CHAPTER 1: INTRODUCTION

This chapter discusses the general background and scope of this research, and concludes with the structure of this thesis.

### 1.1 Facial Expression Detection

Facial expressions are an important channel of nonverbal communication. They convey the emotional states of an individual to observers. Basically, facial muscles produce facial expressions and emotions. These muscles include eye muscles, tongue muscles, internal jaw muscles, lower face muscles, upper face muscles, and so on. These muscles vary in degrees that they contribute to facial expressions, but almost all of the muscles mentioned above can produce visible changes in appearance. Hence the measurement of muscles changes can identify facial expressions, and consequently reveal human emotions.

There are many computer applications that could benefit from facial expression detection. For example, human-computer interaction and computer graphics animation are becoming more frequent nowadays. If computers can capture facial expression instantly, proper process could be taken spontaneously. Thus the barrier between human and the computer will be much reduced.

## 1.2 Edge Maps For Facial Expression Detection

There are many methods for detecting facial expressions; however, these methods are based on analyzing original human facial expression images either manually or with computationally in-efficient algorithms. Original facial images may contain too much data and information for the computer to process fast; hence detecting facial edges can resolve the problem. Edge detection methods can be applied to facial images to produce an edge map of the face. Edge maps can be used to extract structural information of the image. If structural information concerning the face can be obtained from the edge map, then it is possible to use this information for facial expression recognition.

Edge detection is one of the primary means to process an image. Edges in images are areas with strong intensity contrasts, with a jump in intensity from one pixel to the next. The objective of edge detection is to identify a point in a digital image at which pixel intensity changes dramatically. This dramatic change is usually indicated by a point of interest within the image that can be utilized for processing vital events. The purpose of detecting edges in an image is to capture important events. In the ideal case, the result of applying an edge detector to an image may lead to a set of connected curves that indicate the boundaries of objects, the boundaries of surface markings as well as curves that correspond to discontinuities in surface orientation.

Applying an edge detector to a facial image may significantly reduce the amount of data to be processed and may therefore filter out information that may be regarded as less relevant, while preserving the important structural properties of a facial image. If the edge detection process is successful, the subsequent task of interpreting the information contained within the original image may therefore be substantially simplified [1].

### 1.3 Contributions

The purpose of this project is to test an adaptive and efficient bi-dimensional empirical mode decomposition (BEMD) algorithm on facial edge detection, and to develop a metric for facial expression recognition based on the proposed algorithm. The specific contributions include the following:

1. Developed a novel method of using edge maps for facial edge detection.
2. Tested a novel edge detection method (BEMD), refined the parameters for edge detection algorithm.
3. Developed facial feature measurement metrics, and introduced a new metric that is crucial to the measurement system.
4. Tested the proposed algorithm on two major databases to verify the robustness.
5. Developed an identification method to blindly recognize facial expressions.

### 1.4 Structure of the Thesis

This thesis includes four main sections. The literature review section will review the current state of facial expression recognition, edge detection, and the method of empirical mode decomposition. The methodology section will discuss the details of the methods used in the research, including facial edge extraction, the BEMD method, metrics for facial expression recognition the blind detection algorithms that have been developed, and the blind detection algorithm. Two databases were tested and the detection rates will be presented in the results section. The last section will summarize what has been accomplished in this research effort; possible future research directions will also be discussed.

## CHAPTER 2: LITERATURE REVIEW

Since using an edge detector to extract facial edges and recognize facial expressions of human beings is novel, there is no literature that describes this technique. However, facial expression recognition and image edge detection are two mature fields of research. In this chapter, the main techniques used in both fields will be reviewed and the potential of using edge information to recognize facial expression will be revealed. Furthermore, the edge detection method used in this research, BEMD, will be discussed.

### 2.1 Facial Expression Recognition

Facial expressions provide information not only about effective state, but also about cognitive activity, temperament and personality, truthfulness, and psychopathology [2]. Facial recognition techniques detect peoples emotions using their facial expressions. There are many facial recognition techniques that have been developed over the years, mainly in the field of video surveillance, criminal identification, bank card, etc. The most predominant technique is the Facial Action Coding System (FACS), which will be further discussed in the next subsection.

#### 2.1.1 The Facial Action Coding System

FACS [3] is the leading method for measuring facial movement in behavioral science. FACS is currently performed manually by highly trained human experts [2]. It was developed by Ekman and Friesen in 1978.

FACS was developed by determining from palpation, knowledge of anatomy, and videotapes, how the contraction of each of the facial muscles changed the appearance of the face. Ekman and Friesen defined 46 Action Units, or AUs, to correspond to each independent motion of the face. A trained human FACS coder decomposes an observed expression into the specific AUs that produced the expression. FACS is coded from video and the code provides precise specification of the dynamics (duration, onset, and offset time) of facial movement in addition to the morphology (the specific facial actions which occur) [2]. Table 2.1 shows a subset of AUs having been defined.

Table 2.1: Sample subset of AUs

# of AU	Description	Facial muscle	Example image
1	Inner Brow Raiser	Pars medialis	
2	Outer Brow Raiser	Pars lateralis	
4	Brow Lowerer	Corrugator supercili	
27	Mouth Stretch	Digastric	
28	Lip Suck	Orbicularis oris	

Up to now, this system is still the leading method for measuring facial expressions in behavioral science [4]. It has many applications, for example, to demonstrate differences between genuine and simulated pain [5], differences between when people are telling the truth versus lying [6], and differences between the facial signals of suicidal and non-suicidal depressed patients [7]. However, the main drawback of FACS is the time to train human experts and to manually score the video tape. It takes over 100 hours of training to achieve minimal competency on FACS and each minute of video tape takes approximately one hour to score.

Aspects of FACS have been incorporated into computer graphic systems for synthesizing facial expressions, as well as into facial muscle models for parameterizing facial movement [8]. It is important to distinguish FACS itself from facial muscle models that employ aspects of FACS. FACS is performed by human observers using stop-motion video. Although there are clearly defined relationships between FACS and the underlying facial muscles, FACS is an image based method. Facial actions are defined by the image changes they produce in video sequences of face images [2].

### 2.1.2 Facial Expression Measurement

According to [9], Ekman had proposed six basic emotions (happiness, sadness, fear, disgust, anger, and surprise), each of those have a prototypic facial expression, involving changes in facial features in multiple regions of the face, which facilitates analysis. In everyday life, prototypical expressions may occur relatively infrequently, and emotion more often is communicated by changes in one or two discrete features, such as the tightening of the lips, which may communicate anger [10]. A coding example is given to illustrate how FACS works. Figure 2.1 is one of the picture from our reference picture set. When coding experts are scoring AUs of this image, AU 1 (Inner Brow Raiser) and AU 27 (Mouth Stretch) will receive high coding scores. A combination of these two high scores decides the facial expression to be surprise.

A number of methods have been developed for facial recognition include Gabor wavelets [11] [12], linear discriminant analysis [13], local feature analysis [14], and independent component analysis [15]. Recent advances have been made in computer vision for automatic recognition of facial expressions in images. The approaches that have been explored include analysis of facial motion [8], measurements of the shapes of facial features and their spatial arrangements [16], holistic spatial pattern analysis using techniques



Figure 2.1: Recognition example

based on principal component analysis [16], gray level pattern analysis using local spatial filters [17], and methods for relating face images to physical models of the facial skin and musculature [8]. The image analysis techniques in these systems are relevant to the present goals, but the systems themselves are of limited use for behavioral science investigations of the face. Many of these systems were designed with an objective of classifying facial expressions into a few basic categories of emotion, such as happy, sad, or surprised [2].

### 2.1.3 Impact of Cultural Differences On Facial Expression

Cultural differences in intensity ratings of facial expressions by FACS are well documented. One of the most reliable findings has been that Americans tend to pose their expressions more intensely than Japanese [18]. The study was conducted by 107 American and 110 Japanese observers. They viewed Matsumoto and Ekman's (1988) widely used Japanese and Caucasian Facial Expressions of Emotions (JACFEE) set. This set contains 56 slides of seven emotions (anger, contempt, disgust, fear, happiness, sadness, surprise) depicted by eight individuals (four Caucasian and four Japanese; two male and two female

in each group). The expressions were verified as those typically considered universal by FACS.

## 2.2 Edge Detection

Owing to the demand of more efficient and friendly human computer interfaces, research on face image processing has been rapidly growing in recent years. Because people react to multimedia data in the same way as they react to real objects, they react to and behave in front of a computer in the same way as they would interact with another person (except of speaking, which is less frequent in human-computer interaction) [19]. Hence automatic analysis of the users nonverbal behavior conveyed by facial expression which could provide valuable data that the user is currently involved with is our primary means to communicate effectively in human-computer interaction. Once facial expressions are detected with a certain confidence, the system can recognize a facial expression and respond appropriately. Real time recognition and response is required; thus fast processing is needed. Edge information may provide sufficient yet reduced amount of information for detection. Therefore edge detection will be a good method to achieve the goal.

The purpose of edge detection is to discover the information about the shapes and the reflectance or transmittance in an image. It is one of the fundamental steps in image processing, image analysis, image pattern recognition, and computer vision, as well as in human vision.

### 2.2.1 Edge Detection Overview

Currently, a tremendous amount of effort has been made on developing and improving edge detection methods. There are many ways to perform edge detection. The gradient method detects the edges by looking for the maximum and minimum in the first

derivative of the image. This method is incorporated in the Sobel operator [20], the Prewitt operator [21], and so on. Later an improved Sobel operator and algorithm are discussed in terms of optimal threshold in [22]. The Laplacian method searches for zero crossings in the second derivative of the image to find edges. This method is used in the Marr operator [23]. In the late 1980s, the Canny edge detection algorithm started to be known as the optimal edge detector [24] for two dimensional images. This method can give the edge information of both intensity and direction. Deriche [25] introduced a fast recursive implementation of Canny's edge detector, Lanser and Eckstein [26] improved the isotropy of Deriche's recursive filter, and Jahne et al. [27] provide a nonlinear optimization strategy for edge detectors with optimal isotropy. At the same time, The Nalwa-Binford edge detector [28] was chosen to represent the "surface fitting" approach to edge detection.

In the 90s, more edge detectors were published. Sarkar-Boyer edge detector [29] was known as representing the current state of the art in the "zero crossing" approach. Neural networks were also used as a useful tool for edge detection. Since a neural network edge detector is a nonlinear filter, it can have a built-in thresholding capability. Thus the filtering, thresholding operation of edge detection is a natural application for neural network processing. An edge-detection neural network can be trained with back propagation using relatively few training patterns [30]. Another paper [25] investigated the application of neural network for detecting edges in 3-D scales. The results are then used to form the basis for the use of a neural network to carry out automatic edge detection, by defining the correct scale at which to apply, during scale space analysis.

Recently, some other edge detectors have been developed. The representative edge detection methods are improvements based on wavelet transform, mathematical morphology, fuzzy sets, and fractal geometry [31].

### 2.2.2 Predominant Edge Detection Methods

While there are so many edge detection methods available for different research purposes, the widely used and classic methods are described below.

#### Classic Sobel Operator

The Sobel operator is widely used in image processing, particularly within edge detection algorithms. Technically, it is a discrete differentiation operator, computing an approximation of the gradient of the image intensity function. At each point in the image, the result of the Sobel operator is either the corresponding gradient vector or the norm of this vector. Sobel operator is the partial derivative of  $f(x, y)$  as the central computing 3x3 neighborhood at x, y direction. In order to suppress noise, a certain weight is correspondingly increased on the center point, and its digital gradient approximation equations may describe as follows [22]:

$$G_x = \{f(x+1, y-1) + 2f(x+1, y) + f(x+1, y+1)\} \\ - \{f(x-1, y-1) + 2f(x-1, y) + f(x-1, y+1)\} \quad (2.1)$$

$$G_y = \{f(x-1, y+1) + 2f(x, y+1) + f(x+1, y+1)\} \\ - \{f(x-1, y-1) + 2f(x, y-1) + f(x+1, y-1)\} \quad (2.2)$$

Generally, the gradient is:

$$g(x, y) = \sqrt{G_x^2 + G_y^2} \quad (2.3)$$

Or

$$G(x, y) = |G_x| + |G_y| \quad (2.4)$$

Its convolution template operator as follows:

$$T_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \quad T_y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} \quad (2.5)$$

If we use the Sobel operator to detect the edge of image  $M$ , then we can use the horizontal template  $T_x$  and vertical template  $T_y$  to convolute with the image, without taking into account the border conditions, we will obtain two matrices representing the horizontal and vertical discrete partial derivative of the original image. We will denote these two matrices as  $M1$  and  $M2$ . Then the total gradient value  $G$  may be attained by adding the two gradient matrices. Finally, we can get the edge by a thresholding method.

### Canny Operator

Canny proposed a new approach to edge detection that is optimal for step edges contaminated by white noise. There are three common criteria relevant to edge detector performance [24]:

- Good detection. There should be a low probability of failing to mark real edge points, and low probability of falsely marking non-edge points. This criterion corresponds to maximizing signal-to noise ratio.
- Good localization. The points marked as edge points by the operator should be as close as possible to the center of the true edge.
- Only one response to a single edge. It means the single edge has only one response, and the false edge should be mostly restrained.

For the detection criterion, it is defined by the output signal-to-noise ratio (SNR), the larger of the SNR, the higher quality of the edge detection

$$SNR = \frac{|\int_{-W}^{+W} G(-x)f(x)dx|}{n_0\sqrt{\int_{-W}^{+W} f^2(x)dx}} \quad (2.6)$$

where the  $G(x)$  represents the edge function,  $f(x)$  represents the impulse response of the filter of the width of  $W$ . and  $n_0$  represents the mean square deviation of the Gaussian

noise.

For the localization criterion, the positioning accuracy of the edge is defined as follows:

$$Localization = \frac{|\int_{-W}^{+W} G'(-x)f'(x)dx|}{n_0\sqrt{\int_{-W}^{+W} f'^2(x)dx}} \quad (2.7)$$

Where  $G'(x)$  and  $f'(x)$  is the derivative of  $G(x)$  and  $f(x)$ , respectively. The larger the positioning accuracy, the better the result is.

To ensure that the edge has only one response, the average distance  $x_{zc}(f)$  of the zero-crossing point of the derivative of the impulse response of the edge detection algorithm should meet the following criteria:

$$x_{zc}(f) = \pi \left[ \frac{\int_{-\infty}^{\infty} f'(x)dx}{\int_{-\infty}^{\infty} f''(x)dx} \right]^{1/2} \quad (2.8)$$

where  $f''(x)$  is the second derivative of  $f(x)$ .

It can be shown that convolving an image with a symmetric 2D Gaussian and then differentiating in the direction of the gradient form a simple and effective directional operator, which meets the three criteria mentioned above. Suppose  $G(x, y)$  is a 2D Gaussian filter and  $I(x, y)$  is the image, we get the smoothed image  $H(x, y)$  as

$$H(x, y) = G(x, y) * I(x, y) \quad (2.9)$$

and

$$G_n = \frac{\partial G}{\partial n} = \vec{n} \cdot \nabla G \quad (2.10)$$

Ideally,  $\vec{n}$  should be oriented normal to the direction of an edge to be detected, and although this direction is not known a priori, we can form a good estimate of it from the

smoothed gradient direction

$$\vec{n} = \frac{\nabla(G * I)}{|\nabla(G * I)|} \quad (2.11)$$

This turns out to be a very good estimator for edge normal direction for steps, since a smoothed step has strong gradient normal to the edge.

The edge location is then at the local maximum of first derivative of  $H(x, y)$  in the direction  $\vec{n}$ , that is the zero-crossing point of second derivative of  $H(x, y)$

$$\partial^2 H / \partial n^2 = 0 \quad (2.12)$$

This equation illustrates how to find local maxima in the direction perpendicular to the edge. This operation is often referred to as non-maximal suppression (NMS). After NMS operation, the data obtained usually contains some spurious responses. This is called the streaking problem and is quite common in the edge detection problem. These streaking can be eliminated by using a threshold with hysteresis [32].

### Mathematical Morphology Operator

Mathematical morphology theory is developed from geometry. It was introduced by Matheron as a technique for analyzing geometric structure of metallic and geologic samples. It was extended to image analysis by Serra [33]. It is based on set theory; mathematical morphology is established by introducing fundamental operators applied to two sets. One set is said to be processed by another which is known as a structuring element. Let  $I$  denotes a gray-scale two dimensional image,  $B$  denotes the structuring element. The basic mathematical morphological operators are dilation and erosion, derived from these, other operations are also defined.

Mathematical morphology methods can be utilized for image edge detection. Mathematical morphology-based edge detection methods mainly utilize morphological gradient. Similar to the differential gradient operator, the morphological gradient operator can also be combined with a thresholding method to complete the edge detection. The morphological edge detection operator is a nonlinear differential operator, and detected edges are relevant to structural elements. Another application of morphological operations is for binary image processing. They can be used to remove small objects, remove interior pixels, and fills isolated interior pixels, etc. This application is used in our research to refine the edge maps generated by BEMD.

### 2.3 Empirical Mode Decomposition Method

Bi-dimensional Empirical Mode Decomposition (BEMD) is a generalization of the Empirical Mode Decomposition (EMD). It is the key part of the Hilbert-Huang Transform(HHT) which is a new method of time-frequency analysis [34]. This method outweighs other traditional signal analysis when the signals are nonlinear and non-stationary time series. It can identify the instantaneous frequency [35] of the signals at times of interests. With the BEMD method, any signal can be decomposed into a finite and often small number of Bi-dimensional intrinsic mode functions (BIMF) that satisfy the following two conditions: (1) in the whole data set, the number of extrema and the number of zero crossings must either equal or differ at most by one; and (2) at any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero. Through BEMD, any signal can be decomposed into a small number of BIMFs which give meaningful instantaneous frequency. Subsequence application of the HHT to the IMFs will reveal the time-frequency characteristics of the original signal. The BEMD method is adaptive.

HHT/EMD has been further developed and applied to many areas like electric power quality [36], geography [37], medicine [38], finance [39], and so on, including its successful application to image processing. As far as the one-dimensional case is concerned, studies were carried out to show the similarities of EMD with selective filter bank decompositions [40]. Its efficiency for signal denoising was also shown in [41]. These interesting aspects of EMD motivate the extension of this method to Bi-dimensional signals [42]. Lately, a fast and adaptive Bi-dimensional EMD was composed by using order statistic filters to get the upper and lower envelopes [43]. More details of the method used will be discussed in chapter 3.

For BEMD extrema detection and interpolation are carried out using 2D versions of the corresponding 1D method. Let the original image be denoted as  $I$ , a BIMF as  $F$ , and the residue as  $R$ . In the decomposition process  $i^{th}$  BIMF  $F_i$  is obtained from its source image  $S_i$ , where  $S_i$  is a residue image obtained as  $S_i = S_{i-1} - F_{i-1}$  and  $S_1 = I$ . It requires one or more iterations to obtain  $F_i$ , where the Intermediate Temporary State of BIMF (ITS-BIMF) in  $j^{th}$  iteration can be denoted as  $F_{Tj}$ . The steps of the BEMD process can be summarized as below [37],

1. Set  $i=1$ . Set  $S_i = I$ .
2. Set  $j=1$ . Set  $F_{Tj} = S_i$ .
3. Obtain the local maxima map and local minima map of  $F_{Tj}$ .
4. Form the upper envelop and lower envelop of  $F_{Tj}$  by 2D-interpolation of the maxima and minima points.
5. Find the mean envelop as  $M_{Ej} = \frac{U_{Fj} + L_{Ej}}{2}$
6. Calculate  $F_{Tj+1}$  as  $F_{Tj+1} = F_{Tj} - M_{Ej}$ .

7. Check whether  $F_{Tj}$  follows the BIMF properties. These criteria are verified by finding standard deviation(SD) as below

$$SD = \frac{\sum \sum |F_{Tj+1}(x,y) - F_{Tj}(x,y)|^2}{|F_{Tj}(x,y)|^2} \quad (2.13)$$

8. If  $F_{Tj+1}$  meets the criteria given in step 7, take  $F_{Tj} = F_{Tj+1}$ . Set  $i = i + 1$ , and then  $S_i = S_{i-1}$ . Go to the next step. If not, set  $j = j + 1$ , repeat step 3-8 until the stopping criteria is fulfilled.
9. Determine whether  $F_{Tj+1}$  has less than three extreme points, and if so, this is the residue R of image(i.e.  $R = S_i$ ); and the decomposition is complete. Otherwise, repeat step 2-9 to obtain the subsequent BIMFs. In the process of extracting the BIMFs, the number of extrema in  $S_{i+1}$  should be lower than that in  $S_i$ .

The BIMFs and residue of an image together can be named as Bi-dimensional empirical mode components (BEMCs). Except for the truncation error of the digital computer, the summation of all the BEMCs returns the original data/image back as given by

$$\sum_i C_i = I \quad (2.14)$$

where,  $C_i$  is the  $i^{th}$  BEMC.

## CHAPTER 3: METHODOLOGY

In this chapter the detailed methodology of edge extraction and facial expression recognition will be presented, including the illustration of the BEMD method, the development of metrics for expression recognition, procedure of recognition process, and a blind detection algorithm.

### 3.1 Overview of Method

Figure 3.1 shows the flow-chart depicting the overall process of facial expression recognition. One could notice that the process is generally divided into main 2 sub-boxes. The first sub-box indicates the process of edge extraction, including the implementation of BEMD and the use of mathematical morphology. The second sub-box indicates the process of facial expression recognition, including feature extraction and decision making process.

### 3.2 Facial Edge Map Extraction

A more detailed description of edge extraction will be discussed in this subsection, including the implementation of BEMD, mathematical morphology method, and the discussion of important parameters in extracting edges.

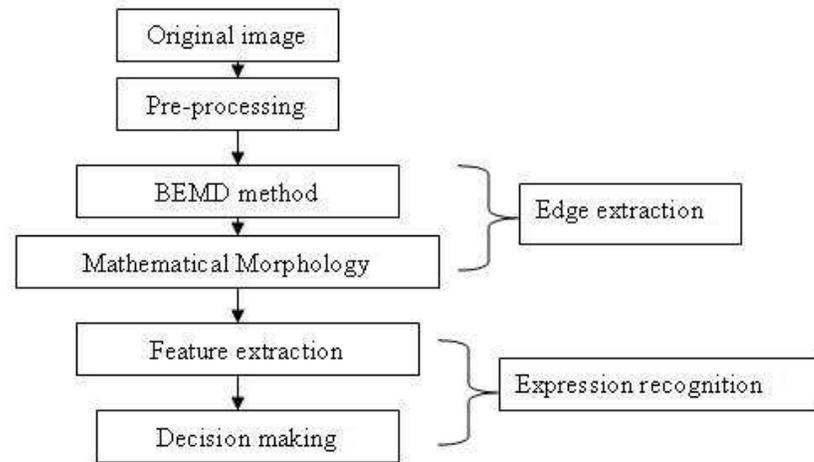


Figure 3.1: Overall process

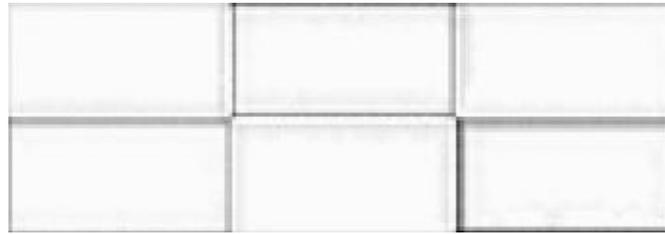
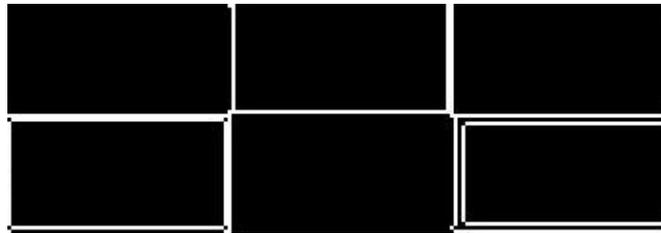
### 3.2.1 BEMD Method

To illustrate how well BEMD method can extract edges, a basic rectangular shape with different levels of grey scale is analyzed, as shown below.

According to the characteristics of the BEMD method, the first IMF/BIMF contains the highest local frequencies. In images, the edges are always with higher frequencies, because they represent sudden intensity changes in an image. Thus in edge detection, the first BIMF is sufficient for extracting important edge information from the original image. Figure 3.2 shows that the 1<sup>st</sup> BIMF extracts those sudden changes of intensity in each rectangular shape very well. Consider the  $512 \times 512$  Lena image in Figure 3.3 for another example. The image is decomposed into six BIMFs and the residue. The original image may be reconstructed by summing each BIMF and adding the residue to the result. It is obvious that the 1<sup>st</sup> BIMF did a good job in extracting the edge.



(1)Original picture

(2)The 1<sup>st</sup> BIMF

(3) extracted edge image

Figure 3.2: Examples using basic rectangular shape

### 3.2.2 Mathematical Morphology

The edge detection results of the Lena image are presented below in Figure 3.4 by converting a gray-scale image of the first BIMF to a binary image. A mathematical morphology method is also implemented to remove small areas smaller than a suitable critical value from the edge image. The implementation of the morphology method requires the selection of a critical value for determining which areas are removed and which are left

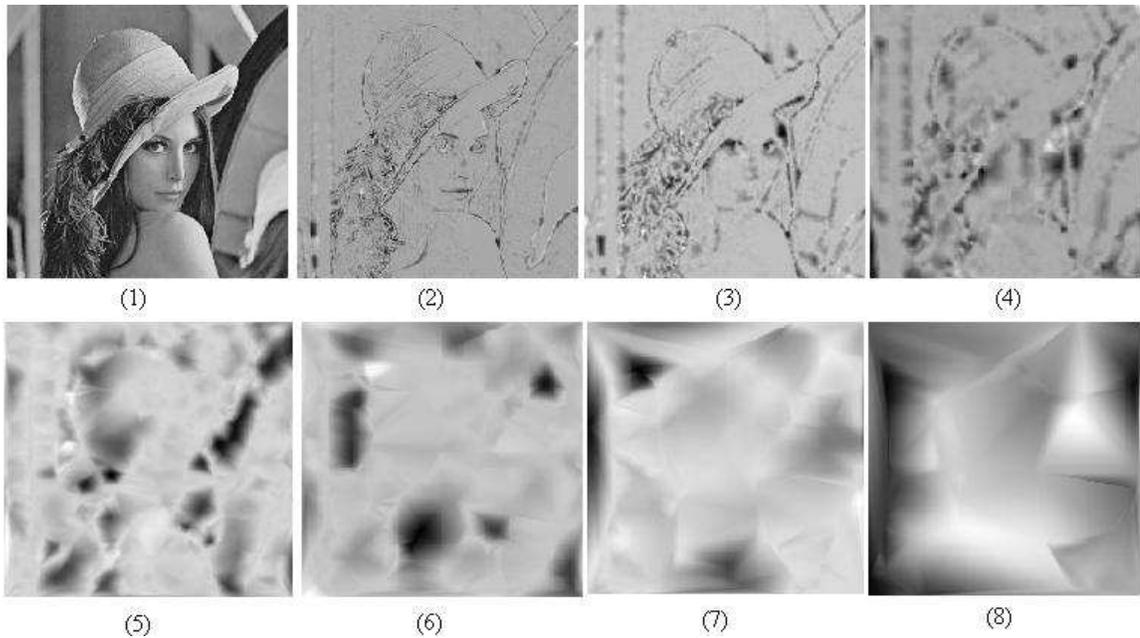


Figure 3.3: BEMD of Lena image: (1) Original image, (2)1<sup>st</sup> IMF, (3)2<sup>nd</sup> IMF, (4)3<sup>rd</sup> IMF, (5)4<sup>th</sup> IMF, (6)5<sup>th</sup> IMF, (7)6<sup>th</sup> IMF, (8)Residue

unchanged. The critical value is equal to a count of contiguous pixels in a given feature of the edge image. If the critical value is too low, then too many unimportant features will remain in the edge image. If the critical value is too high, then too many important features of the edge image will be removed.

The results are compared with different critical values (from 0 pixels to 80 pixels) and determined 40 pixels as the optimal critical value which can obtain the best edge image without losing important edge information. Selected results with different critical value are shown in Figure 3.5.

Figure 3.5 shows too many details were retained when using 20 pixels, while too many details were removed when 60 pixels were used. A balanced value of 40 pixels provides the most satisfactory results.

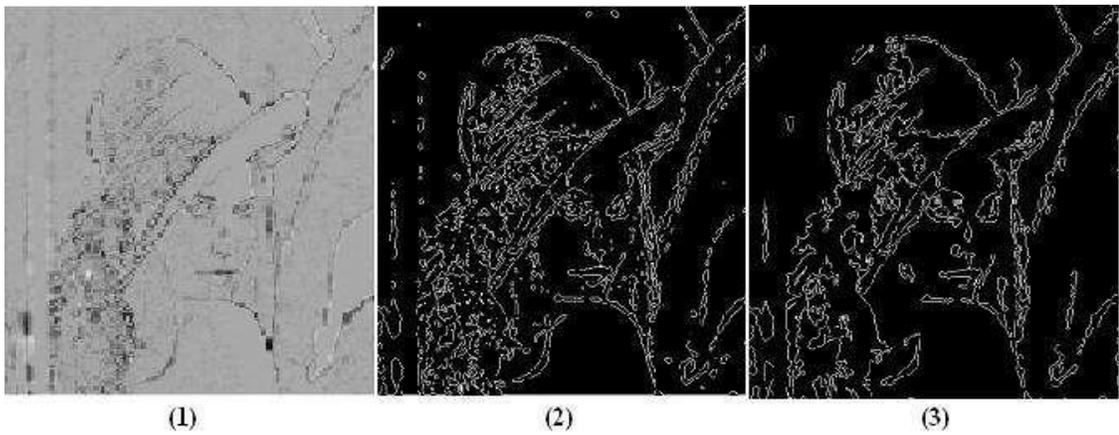


Figure 3.4: (1) The 1<sup>st</sup> BIMF of Lena image (2) Converting the grey scale image to binary image (3) Edge results after applying mathematical morphology

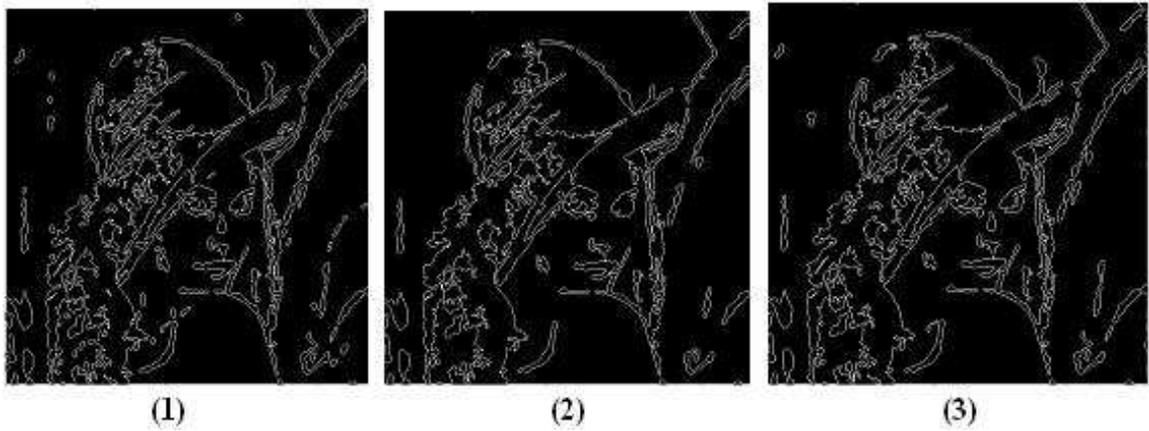


Figure 3.5: Edge images obtained with different critical values in morphology operation: (1) 20 pixels (2) 60 pixels (3) 40 pixels

### 3.3 Facial Expression Recognition Using Edge Map

Generally, facial expressions are determined by extraction of certain facial features. They are divided into two categories based on those facial features: category I includes "Happy", "Sad", and "Fear", and category II includes "Surprise", "Angry", and "Disgust". In this section, the establishment of facial expression recognition criteria will be discussed and recognition procedure will be proposed.

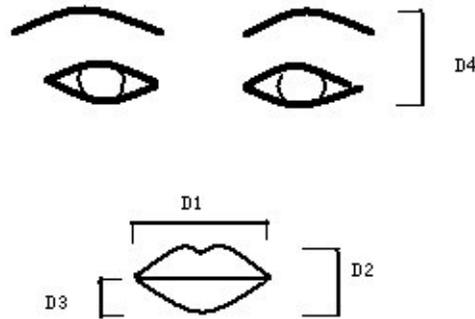


Figure 3.6: Facial skeleton and characteristic distances

### 3.3.1 Facial Expression Recognition Criteria

In this research, three predominant features (D1, D2, and D4) and 1 new feature (D3) were included for facial expression detection. These four features are distance measures as specified in Table 3.1. Figure 3.6 also shows the facial skeleton with which the metrics were defined. The reason why D3 is important to the measuring system lies in that it is an important parameter to distinguish between "Happy" and "Sad". Therefore, I added D3 as the 4<sup>th</sup> metric.

Table 3.1: Four facial features for facial recognition

D1	Mouth opening width, distance between right and left corners
D2	Mouth opening height, distance between upper and lower lips
D3	Distance between mouth corners and lower lip
D4	Distance between eyebrow and lower eyelid

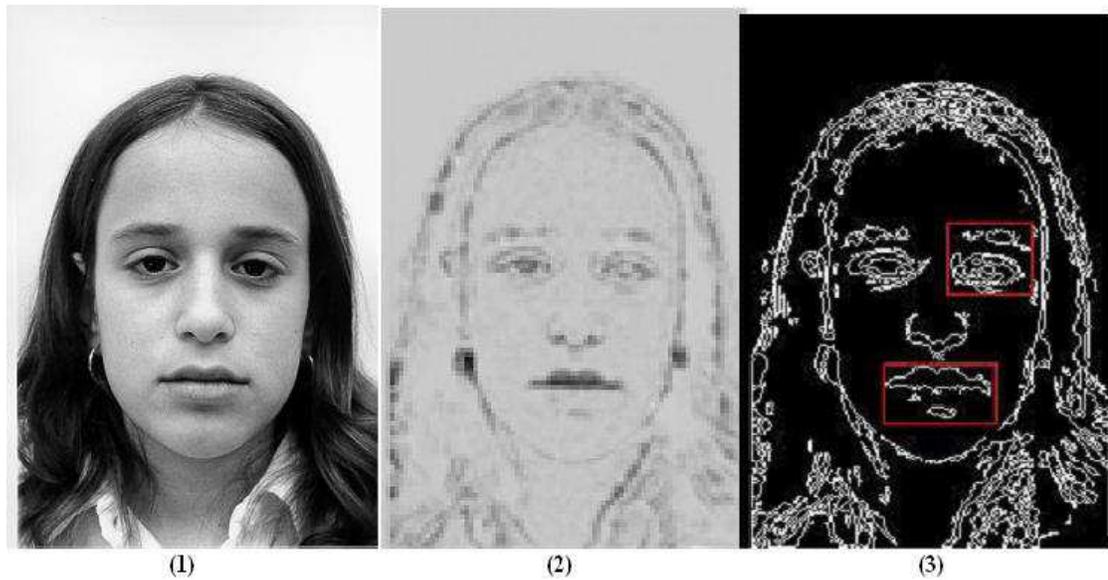


Figure 3.7: Neutral facial expression: (1) Reconstructed image from BIMFs, (2) The 1<sup>st</sup> BIMF, (3) Extracted edge image with bounding boxes

### 3.3.2 Determination of Facial Expression Recognition Criteria

To develop the metrics for facial expression recognition, an American models picture set which is developed by a psychologist was chosen as the reference. The dataset consists of seven images representing "Neutral", "Happy", "Sad", "Fear", "Angry", "Disgust", and "Surprise".

The "Neutral" expression in this picture set was used as a reference picture for other facial expressions. D1-D4 were measured on the edge map extracted from the 1<sup>st</sup> BIMF of "Neutral" expression as shown in Figure 3.7. A bounding box is manually placed over the mouth and the left eye of the edge image for the purpose of computing the distance measures, D1 through D4. The measurements are in units of pixels.

Figures 3.8 through 3.13 show the original reference pictures and extracted edge maps of other six basic expressions. D1-D3 were then measured on the edge maps of these edge maps. D4 is only needed for "Surprise" and "Disgust" as shown in Figure 3.6.

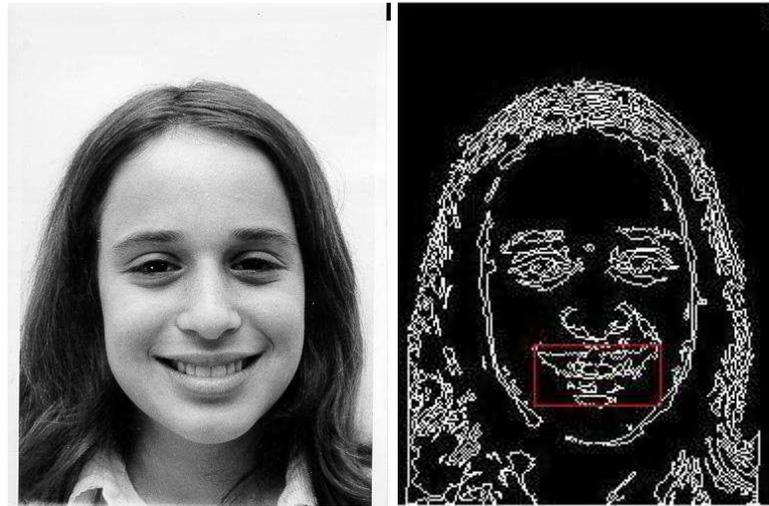


Figure 3.8: "Happy" and edge map for recognition

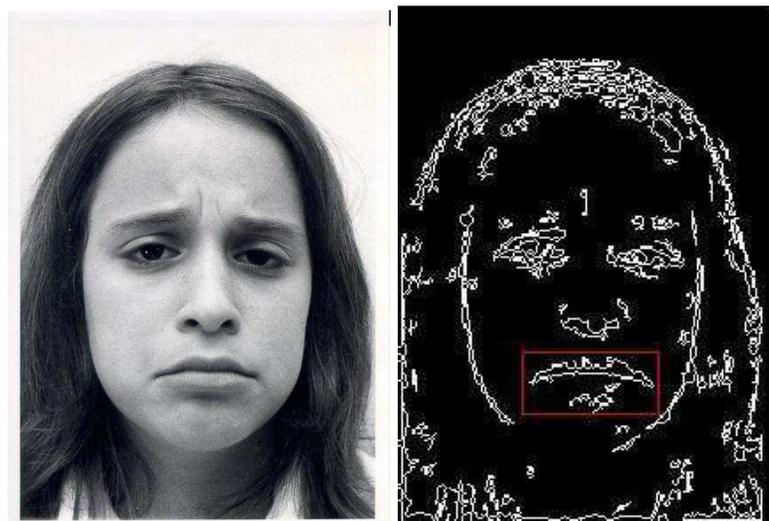


Figure 3.9: "Sad" and edge map for recognition

Table 3.2 shows the decision metrics for facial expressions. In each cell, the value shows the number of pixels measured for each metric, and the percentages are the percent-change of each metric from its corresponding "Neutral" expression. It can be observed from the table that an increase in D1 puts expressions in category I (C-I) and a decrease in category II (C-II) In C-I, a significant increase in D3 separate "Happy", a significant



Figure 3.10: "Fear" and edge map for recognition

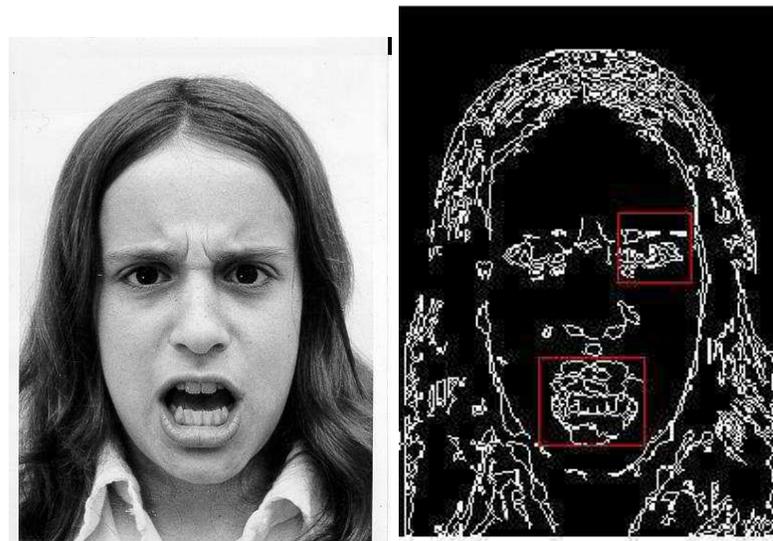


Figure 3.11: "Angry" and edge map for recognition

decrease for "Sad", and the rest are "Fear". In C-II, a decrease in D4 recognizes "Angry", a large change of both D2 and D4 identifies "Surprise", and an insignificant change in D2 and D4 specifies the expression "Disgust".

As having been introduced in previous chapter, there are a lot of classic edge de-

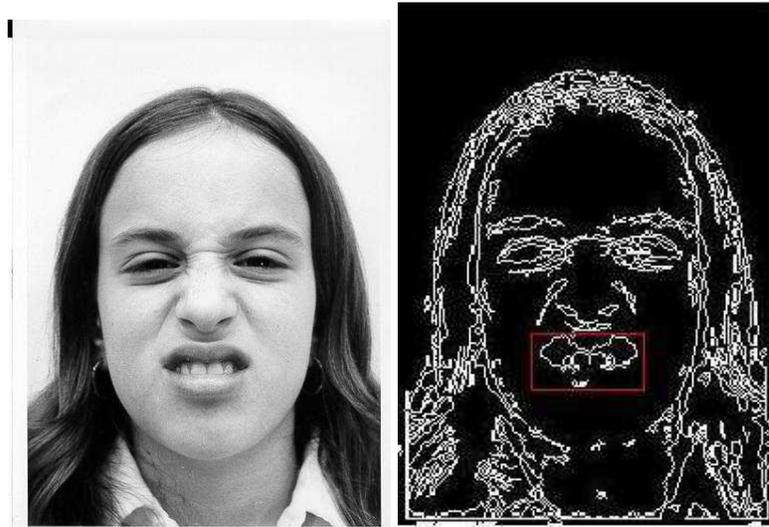


Figure 3.12: "Disgust" and edge map for recognition

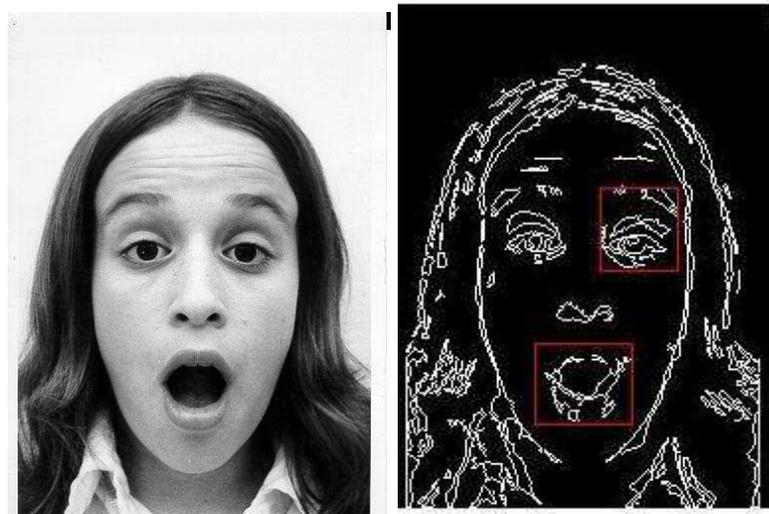


Figure 3.13: "Surprise" and edge map for recognition

tection method have been developed so far. However, for facial expression, using BEMD method to extract the edges is superior to other predominant edge detection methods such as Canny and Sobel. That is because the BEMD method can extract important edges which are used for measuring without being interfered by surrounding unimportant edges

Table 3.2: Criteria for expression recognition

	Happy	Sad	Fear	Neutral	Surprise	Angry	Disgust
D1	61(20%)	60(18%)	61(20%)	51(0%)	43(-16%)	45(-12%)	46(-10%)
D2	/	/	/	24(0%)	40(40%)	36(50%)	29(21%)
D3	25(79%)	11(-21%)	21(50%)	14(0%)	/	/	/
D4	/	/	/	29(0%)	44(52%)	24(-17%)	30(3%)

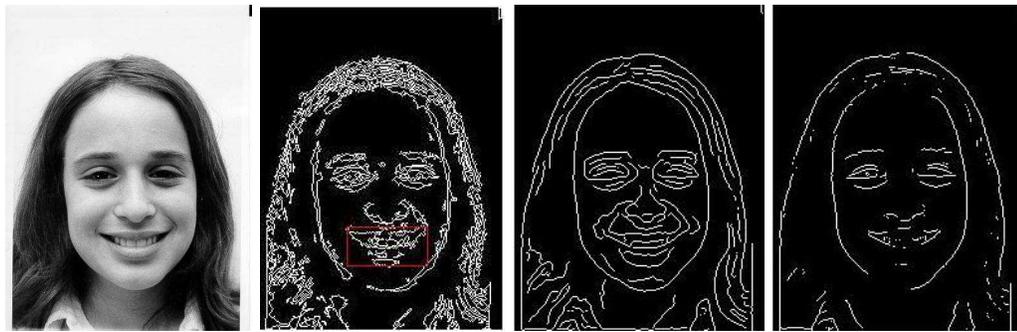
or losing too many essential edges.

A comparison study was conducted to compare the quality of extracted edge images by BEMD, Canny, and Sobel. The two example results from "Happy" and "Fear" expressions are shown in figure 3.14 reveal that the Canny method extracted too many details which makes it hard to measure specific metrics, while the Sobel method did not extract enough edge details which could prevent the algorithm from accurately measuring facial features.

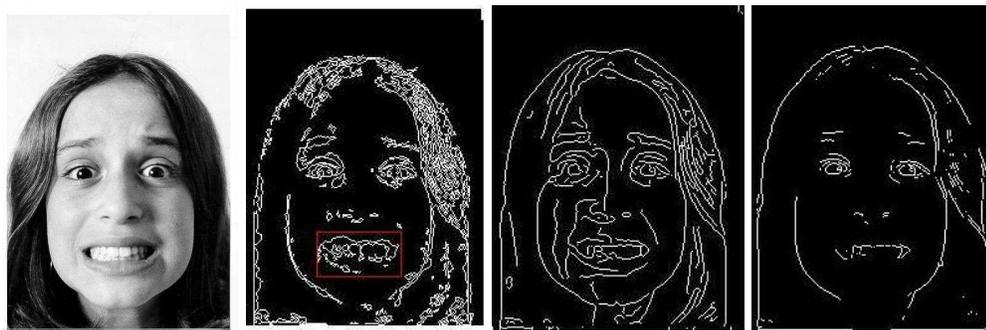
### 3.3.3 Facial Expression Recognition Procedures

To sum up the facial expression recognition procedures, the procedure of developing facial expression recognition criteria is summarized as follows:

1. A reference dataset was chosen. The dataset consists of seven images: A "Neutral" face and six typical facial expressions.
2. D1-D4 were measured on the edge map of the "Neutral" face, these measurements was used as references.
3. D1-D3 of each expression were measured and D4 was only needed for the expressions of "Angry", "Surprise" and "Disgust". These data was then converted as percentage changes from the corresponding "Neutral" measurements.



(1) Happy (2) BEMD method (3) Canny method (4) Sobel method



(1) Fear (2) BEMD method (3) Canny method (4) Sobel method

Figure 3.14: Comparison with predominant edge detection methods

Therefore the facial expression recognition procedures were developed based on the above criteria. These procedures are summarized below and are presented graphically in Figure 3.15.

1. D1 determines if the expression is in C-I or C-II.
2. For C-I expressions, D3 is used to distinguish "Happy", "Sad", and "Fear".
3. For C-II expressions, D2 and D4 are used to recognize "Angry", "Surprise", and "Disgust".

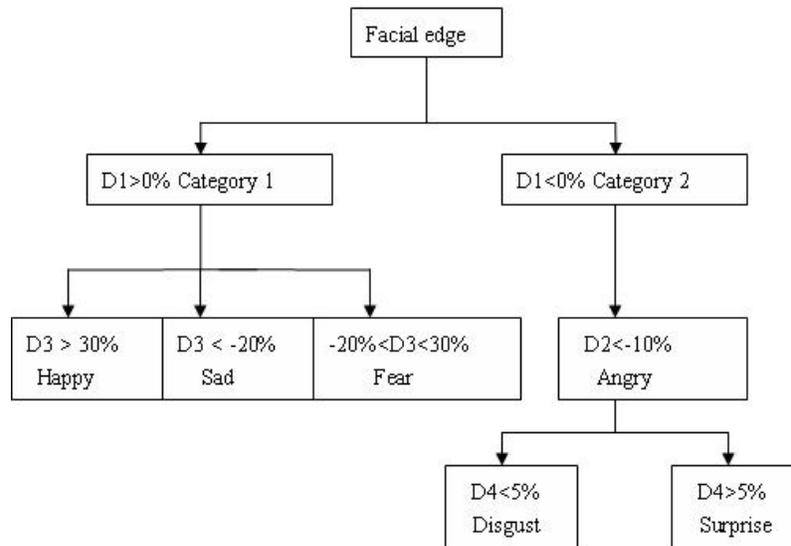


Figure 3.15: Procedures and thresholds for detection

### 3.4 Facial Expression Blind Detection

In previous work, the measuring boxes in edge maps were cropped manually with human's observation, which still took more time than blind recognition. In this research, a blind detection algorithm was developed to test the detection accuracy of facial expressions. Initial results suggest great potential of the proposed method for blind-detection that may lead to even more efficient techniques for facial expression recognition. Figure 3.16 shows the flow chart of the blind detection process. The strategy is firstly to roughly establish two boxes according to the position of eye area and mouth area on the extracted edge map of the neutral expression, and then expand the boxes by pixels from each of the 4 directions until the number of pixels calculated in the boxes will not change. Therefore, the boxes which can crop the target area were developed. An example of how this strategy works for the mouth area is shown in figure 3.16. Initially the algorithm creates a rough estimate of the mouth area as depicted in the red box. Then the sides expand outward until there is no increase in the number of bounded pixels. The final bounding box is shown by

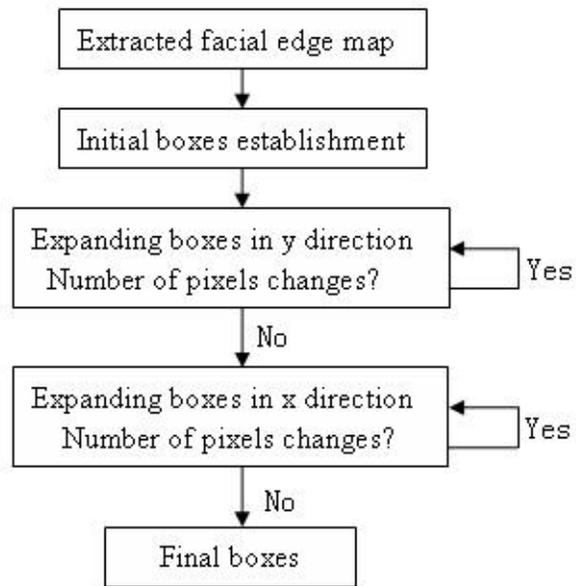


Figure 3.16: Strategy examples for blind detection for mouth area

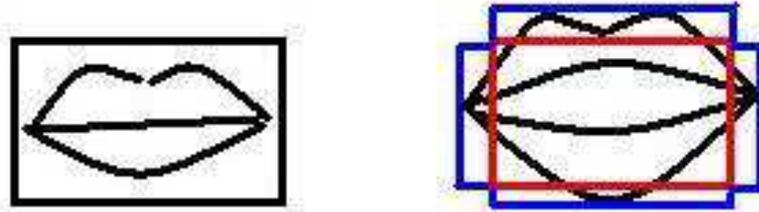


Figure 3.17: Strategy examples for blind detection for mouth area

the blue area in figure 3.17.

## CHAPTER 4: RESULTS

This chapter will present facial expression detection results as well as the results from blind detection by using the Japanese Female Facial Expression (JAFFE) and the Taiwanese Facial Expression Image (TFEID) databases, the two predominant image databases for analyzing facial expressions. The accuracy of the detection results will be presented in terms of the categorical and success detection rates for each expression. Cultural differences that affect facial expression detection will also be discussed.

### 4.1 The Japanese Female Facial Expression (JAFFE) Database

The JAFFE database contains 213 images of seven facial expressions (6 basic facial expressions + 1 neutral) posed by ten Japanese female models. The database was planned and assembled by Miyuki Kamachi, Michael Lyons, and Jiro Gyoba [12]. The photos were taken in the Psychology Department at Kyushu University. In this research, It was found that there was one model whose expressions were very difficult to discern even with the naked eye. So this model's image set was deleted and in total 192 images by nine models were tested.

As discussed in Chapter 3, the six facial expressions are firstly divided into two categories , one category for "Happy, "Sad" and "Fear" and a second category for "Angry", "Surprise" and "Disgust". The metric D1 was applied to the two data bases to distinguish category I from category II by computing the percentage change in D1 from that of the neutral image. Figure 4.1 shows categorical detection results using D1. D1T is the thresh-

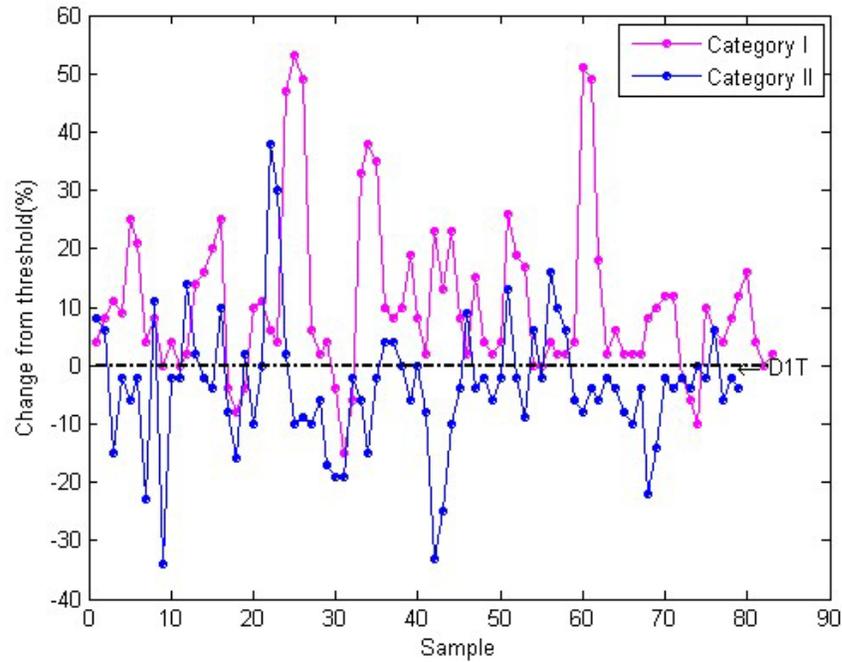


Figure 4.1: Categorical detection results for JAFFE

old of D1 determined by the "Neutral" expression. A value of D1 above D1T decides the expressions in Category-I (C-I). In other words, points above D1T in Figure 4.1 are determined as C-I, otherwise, the expression belongs to Category-II (C-II). The success rates were 89% for C-I and 78% for C-II for the JAFFE database.

Table 4.1: Categorical success rate for JAFFE

	Category I	Category II
Detection rate	89%	78%

Figure 4.2 shows the results for subcategorizing C-I by applying D3 to 192 images selected from the JAFFE database. D3H and D3S refer to D3 thresholds for "Happy" and "Sad", respectively. The D3 values that fall above D3H belong to "Happy", D3 values

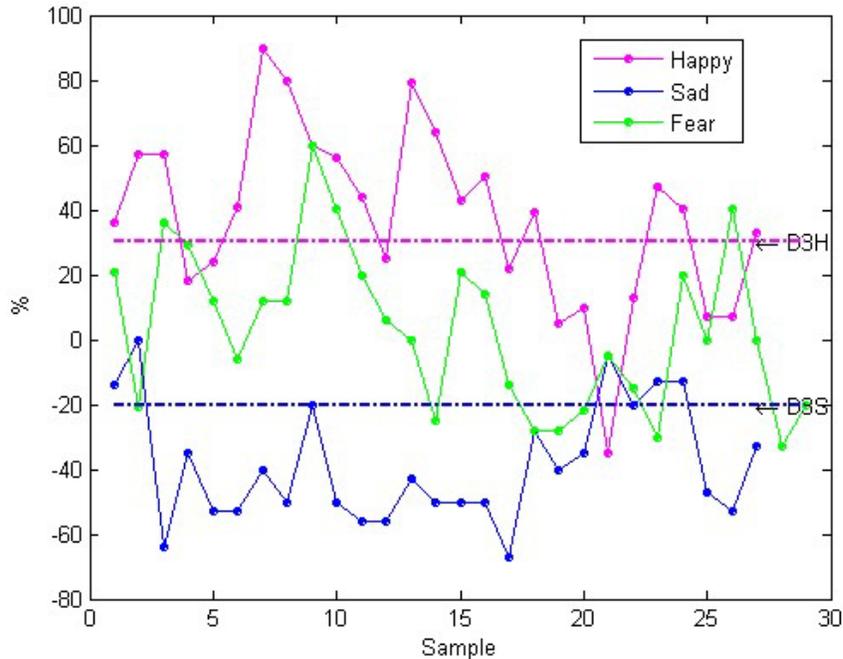


Figure 4.2: Detection results for "Happy", "Sad", "Fear"

that fall below D3S belong to "Sad", any D3 values that fall between D3H and D3S are classified as "Fear". It can be seen from Figure 4.2 that each of these three expressions are well scattered by D3 thresholds, with good detection rates of 63%, 81%, and 67%, respectively.

In C-II detection, it was found that facial expressions of certain emotions of the tested databases are different from the American models reference set in Chapter 3. This is especially true for C-II expressions. Table 4.2 shows the percentage changes in D2 and D4 on the subjects in JAFFE databases under test. As can be observed, D2 for "Angry" has a significant negative change from "Neutral". Hence, D2 was used to identify "Angry" first. A combination of D2 and D4 are then used to separate "Surprise" and "Disgust" as discussed in the previous chapter. Once this adjustment was made, the adjusted metrics

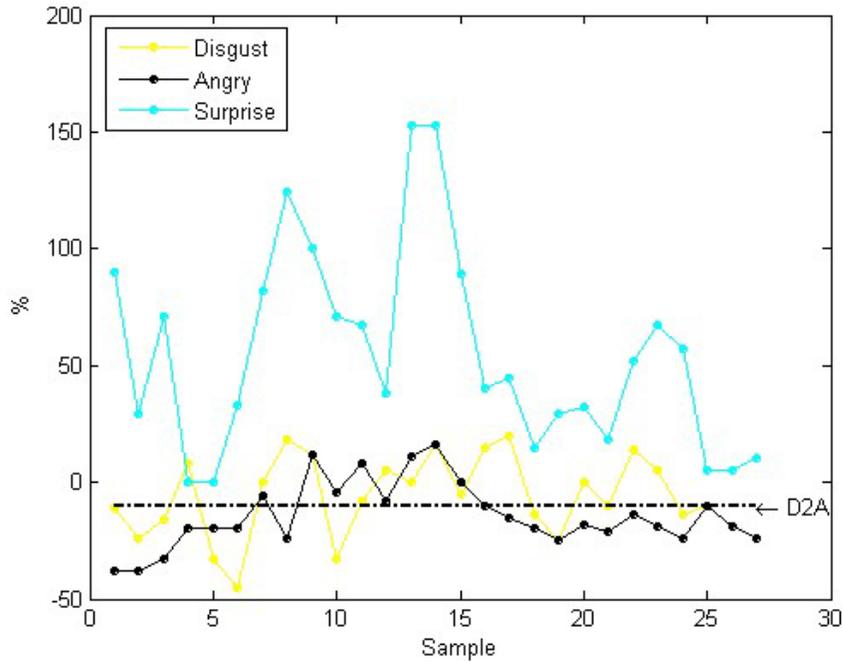


Figure 4.3: Detection results for "Angry" of JAFFE

were applied in detecting expressions for all subjects in both databases.

Table 4.2: Percentage changes of D2 and D4 in JAFFE

	Neutral	Angry	Disgust	Surprise
D2	0%	-14%	-4%	55%
D4	0%	-16%	-9%	18%

The results for facial expressions detection in C-II are shown in Figures 4.3 and 4.4. D2A is the threshold to recognize "Angry" from "Surprise" and "Disgust". As demonstrated from Figure 4.3 "Angry" has the most significant negative change from "Neutral" in D2. A -10% threshold was set to recognize "Angry", D2 values that fall below D2A are recognized as "Angry".

Finally, "Surprise" and "Disgust" were separated using D4. As shown in Table

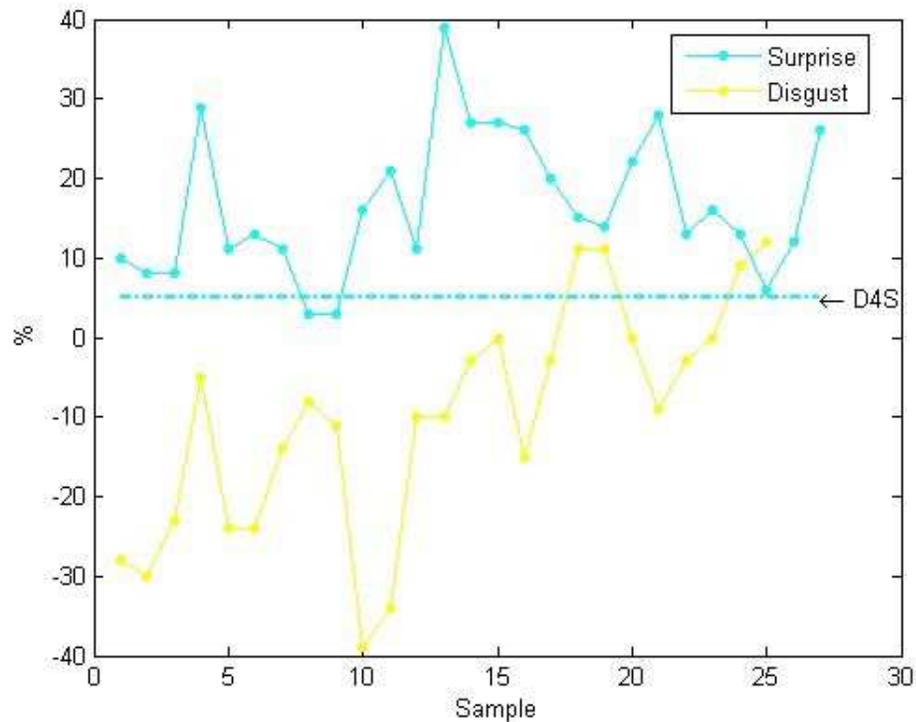


Figure 4.4: Detection results for "Surprise" and "Disgust" of JAFFE

4.2, it shows a decrease in D4 for "Surprise" but an increase in D4 for "Disgust". A 5% threshold was set to distinguish these two groups, where a change greater than 5% identifies "Surprise" while a change less than 5% indicates "Disgust". As shown in Figure 4.4, any D4 value that falls above is recognized as "Surprise", otherwise are recognized as "Disgust". Table 4.3 below shows the detection rate for each individual facial expression for the JAFFE database. The raw data are listed in the appendix. The detection rates are the percentages of each expression having been successfully recognized out of the total number of images to be recognized.

One could notice that none of the rate is below 60% for JAFFE database, which indicate the feasibility of my proposed method. The method will be further analyzed by testing on another predominant database.

Table 4.3: Detection rate for 6 facial expressions of JAFFE

	Category I			Category II		
	Happy	Sad	Fear	Disgust	Angry	Surprise
Detection rate	63%	81%	67%	68%	70%	93%

## 4.2 Taiwanese Facial Expression Image Database (TFEID)

The TFEID consists of 7200 face images captured from 40 models (20 males), each with eight facial expressions: "Neutral", "Angry", "Contempt", "Disgust", "Fear", "Happy", "Sad" and "Surprise". Models were asked to gaze at two different angles (0 and 45 degrees). Each expression includes two kinds of intensities (high and slight) and was captured by two CCD-cameras simultaneously with two viewing angles (0 and 45 degrees) [44]. In this research, 280 images which include seven front view basic expressions of forty models out of the whole database were used.

The same criteria and detection process was applied to TFEID database as those used for the JAFFE database. Figure 4.5 shows the categorical detection results for images selected from the TFEID database. Points above D1T in Figure 4.5 are determined as C-I, otherwise, the expression belongs to C-II. Table 4.4 shows the categorical detection rates for these same test cases. The categorical detection rate is the percentage of the expressions which have been separated to the right category out of original pool.

Table 4.4: Categorical success rate for TFEID

	Category I	Category II
Detection rate	93%	79%

Similarly, D3 was used to distinguish "Happy", "Sad", and "Fear". D3H and D3S refer to D3 thresholds for "Happy" and "Sad", respectively. Any D3 values that fall be-

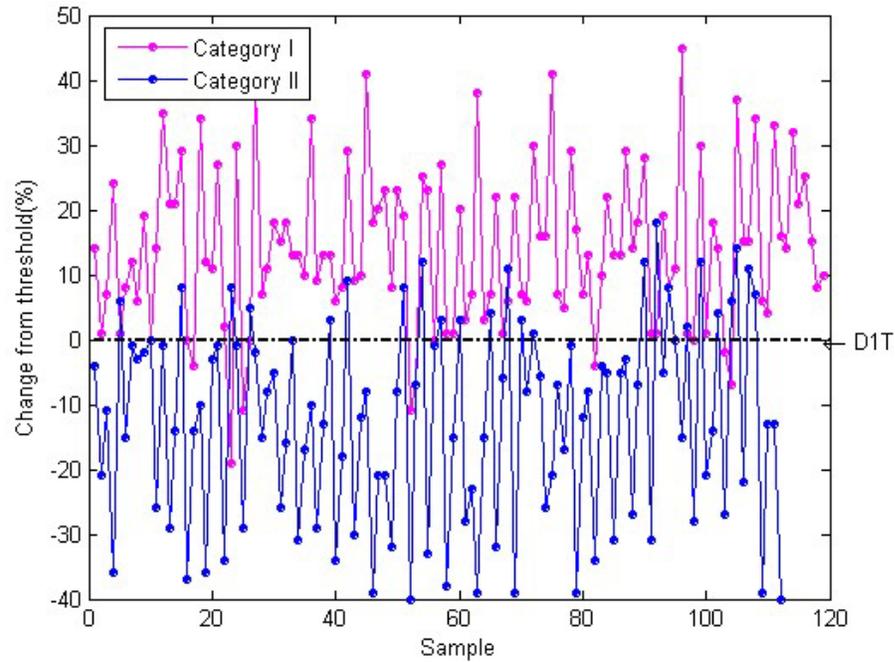


Figure 4.5: Categorical detection results for TFEID

tween D3H and D3S are classified as "Fear". Figure 4.6 shows the detection results for C-I for TFEID. As demonstrated from Figures 4.7 and 4.8, D2A is the threshold to recognize "Angry" from "Surprise" and "Disgust". Lastly, D4 is used for separating "Disgust" and "Surprise". Table 4.5 summarizes the detection rates of six expressions for TFEID. The raw data are listed in the appendix.

Table 4.5: Detection rate for 6 facial expressions of TFEID

	Category I			Category II		
	Happy	Sad	Fear	Disgust	Angry	Surprise
Detection rate	90%	76%	73%	51%	82%	97%

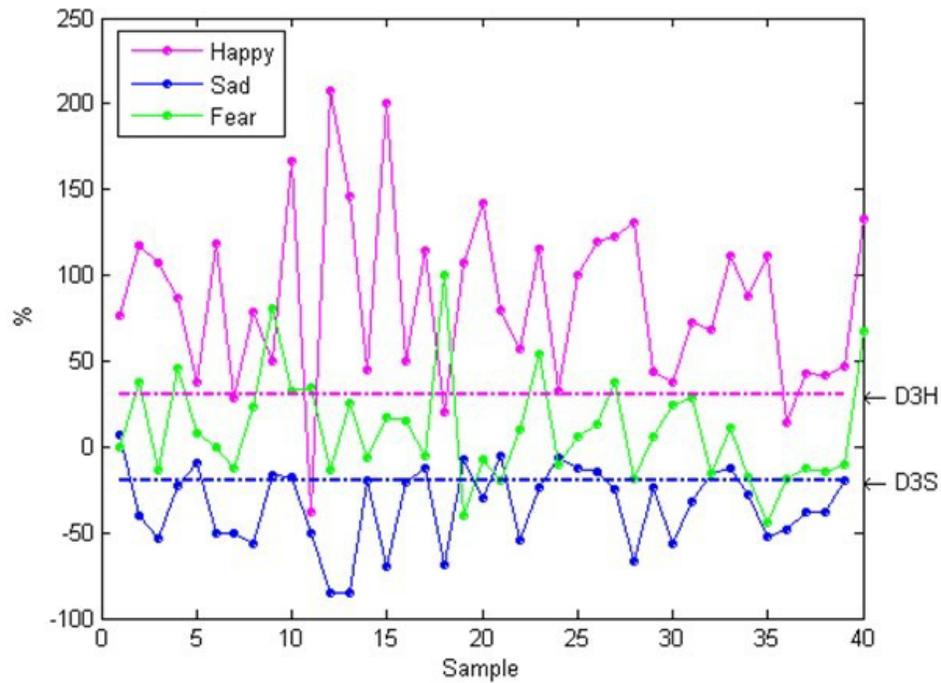


Figure 4.6: Detection results for "Happy", "Sad", and "Fear" of TFEID

### 4.3 Culture Differences and Facial Expressions

By comparing the detection rates from JAFFE in Table 4.3 with the detection rates from TFEID in Table 4.5, one will notice that the detection algorithm generally performed better on the TFEID database. The reason is that the forty models were well guided by physiologists on how they should pose their expressions while the Japanese models were not. This made the expressions easier to recognize for the TFEID.

While detection rates are relatively high for both databases in all categories, one can easily observe that detection of "Disgust" is relatively low for both databases. As mentioned in Chapter 2, according to [18], "cultural differences occurred on multiple emotion scales for each expression" This may be due to the differences between the ways of expressing emotions of different cultural groups. Figure 4.9 shows four examples of the

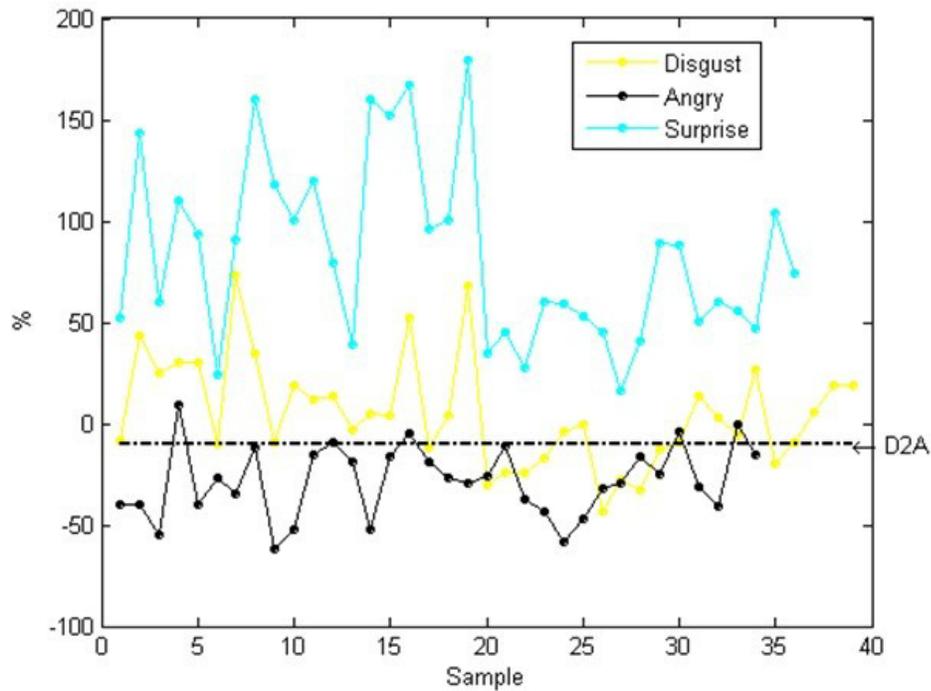


Figure 4.7: Detection results for "Angry" of TFEID

same model from the JAFFE database whose detection results could be affected by the way of posing expressions. In contrast with American model reference set in Chapter 3, Table 4.6 summarizes D1-D4 values compared with "Neutral" of this Japanese model. It shows how culture difference could affect the recognition results.

Table 4.6: D1-D4 values of one of the Japanese model

	Happy	Sad	Fear	Neutral	Surprise	Angry	Disgust
D1	67(26%)	55(4%)	53(0%)	53(0%)	52(-2%)	52(-2%)	51(-4)
D2	/	/	/	20(0%)	28(40%)	18(-10%)	19(-5%)
D3	27(50%)	11(-21%)	13(-28%)	18(0%)	/	/	/
D4	/	/	/	34(0%)	43(27%)	27(-21%)	27(-21%)

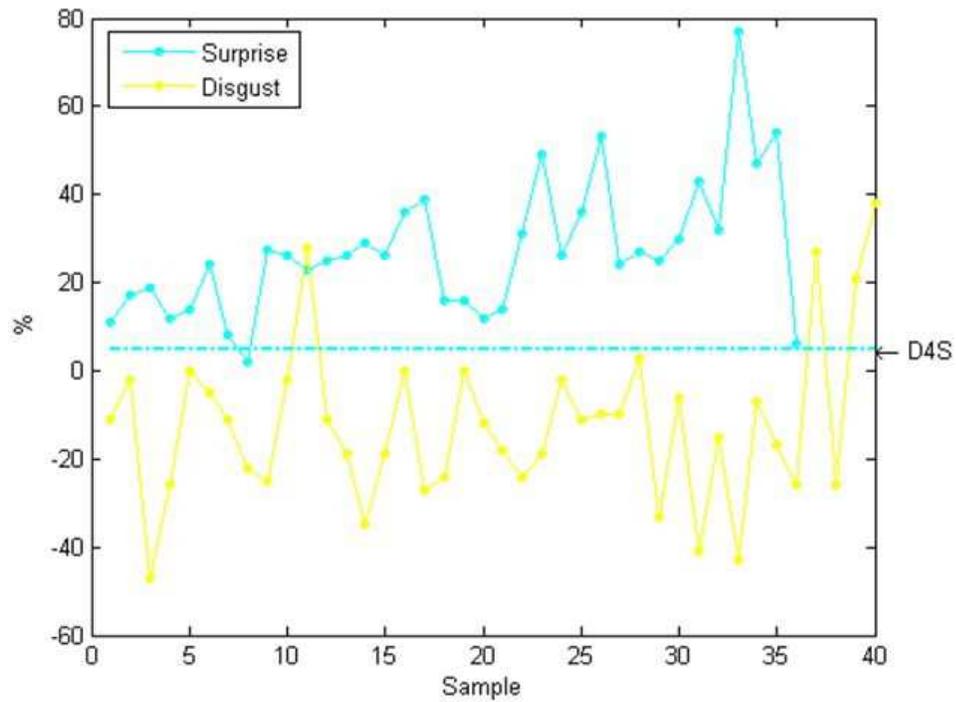


Figure 4.8: Detection results for "Disgust" and "Surprise" of TFEID

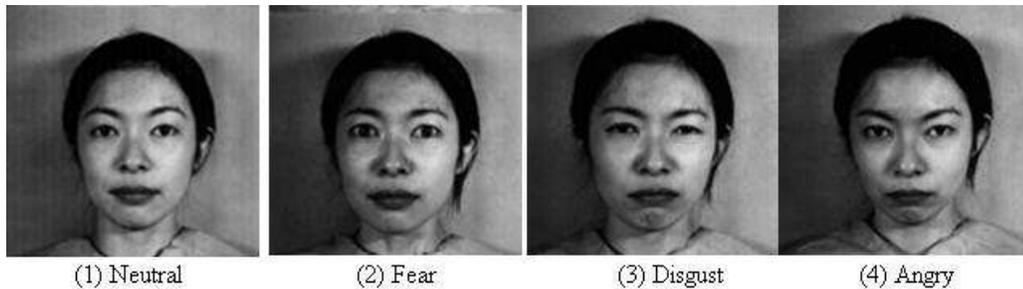


Figure 4.9: Japanese Female Facial Expressions

#### 4.4 Blind Detection For Both of The Databases

As discussed in Chapter 3, detection methods used in this research effort is based on manually cropping interested areas by observation. This section will present the detection results tested on both of the databases by employing a blind-detection algorithm.

By testing upon the JAFFE and TFEID databases, it was found that due to the

robustness of the blind detection algorithm, the initially set boxes of interest areas are the same for the two databases. That will cause some misjudgment for the program to stop before the box has found the whole area. Consequently a result of this effect is that the D3 measure tended to be smaller using blind detection in comparison with using the manual method for selecting interest areas. Thus the threshold of D3H and D3S was changed to 20% and -15%, respectively. Table 4.7 through Table 4.10 show the success detection rates for both of the databases using these new threshold. The raw data will be listed in the appendix.

Table 4.7: Categorical success rate of the blind-detection algorithm tested on the JAFFE

	Category I	Category II
Detection rate	78%	62%

Table 4.8: Detection rate for 6 facial expressions of the blind-detection algorithm tested on the JAFFE

	Category I			Category II		
	Happy	Sad	Fear	Disgust	Angry	Surprise
Detection rate	63%	67%	55%	44%	56%	78%

Table 4.9: Categorical success rate of the blind-detection algorithm tested on the TFEID

	Category I	Category II
Detection rate	87%	65%

Table 4.10: Detection rate for 6 facial expressions of the blind-detection algorithm tested on the TFEID

	Category I			Category II		
	Happy	Sad	Fear	Disgust	Angry	Surprise
Detection rate	68%	41%	45%	44%	54%	89%

#### 4.5 Summary of Results

To sum up, a set of recognition criteria was developed through previous literature and experiments. A new metric was added to increase recognition success rates. To show the robustness of this method, criteria were established by expressions of an American model, and databases from two different (yet related) cultural groups. The same metrics were applied to the images from both databases. Due to limitation of accessible and valid database source, BEMD-detectable Caucasian facial expression database was not available at this moment, thus only Asian facial expression databases were tested in this research. Those two databases are widely used facial expression image databases.

From the above results, it reveals that the computational method based on BEMD is efficient and capable of facial expression recognition comparing with other facial expression whose detection rates are generally around 70%. The blind detection results also indicated the robustness of this method, although future research still needs to be done to improve the detection rates.

## CHAPTER 5: CONCLUSIONS AND FUTURE WORK

This chapter will restate the research objective, conclude what have been accomplished in the research presented in this thesis, and summarize possible future research directions.

### 5.1 Restatement of the Problem and Objectives

Traditional facial expression recognition methods are based on direct examination of original pictures. They require subjective expertise and are time consuming. Thus to apply an edge detector to a facial image may significantly reduce the amount of data to be processed by the computer. BEMD method is one of the efficient edge detection methods, because image signals are nonlinear and non-stationary signals and BEMD is especially efficient for analyzing these kind of signals.

Since there is no known edge-based detection method for facial expression recognition, in this research our goal is to develop a metric which can effectively recognize facial expressions by using BEMD edge detection method, consequently potentially eliminate human observation of images.

### 5.2 Conclusions

In this thesis, the contributions are concluded as follows:

- A BEMD based edge detection algorithm was developed, facial expression measurement metric were created, and intensive database testing was conducted.
- The success rates of recognition suggest that the proposed method could be a potential alternative to traditional methods for human facial expression recognition with minimal human interactions.
- Initial results suggest great potential of the proposed method for blind-detection that may lead to even more efficient techniques for autonomous facial expression recognition.

### 5.3 Future Work

Although overall results are encouraging, at the same time, it is also suggested by the results that further research needs to be done to further improve the detection rates. One potential solution may be to add yet another metric to strengthen the recognition rates. A more efficient and accurate blind detection algorithm will be needed to improve the recognition rates as well. Meanwhile, this topic could be extended to possible real-time detection which leads to even more efficient facial expression recognition technology.

## BIBLIOGRAPHY

- [1] L. F. Liang and Z. L. Ping, “An edge detection algorithm of image based on empirical mode decomposition,” *Second International Symposium on Intelligent Information Technology Application*, vol. 1, pp. 128–132, 2008.
- [2] G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski, “Classifying facial actions,” *Pattern Analysis and Machine Intelligence*, vol. 21, no. 10, pp. 974–989, 1999.
- [3] P. Ekman and W. Friesen, *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto, Calif: Consulting Psychologists Press, 1978.
- [4] P. Ekman and E. L. Rosenberg, *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression using the Facial Action Coding System (FACS)*. New York: Oxford Univ. Press, 1997.
- [5] K. D. Craig, S. A. Hyde, and C. J. Patrick, “Genuine, suppressed, and faked facial behavior during exacerbation of chronic low back pain,” *Pain*, vol. 46, pp. 161–172, 1991.
- [6] P. Ekman, “Telling lies: Clues to deceit in the marketplace,” in *Politics and Marriage*, New York: W.W. Norton, 1985.

- [7] M. Heller and V. Haynal, "The faces of suicidal depression (translation les visages de la depression de suicide)," *Kahiers Psychiatriques Genevois*, vol. 16, pp. 107–117, 1994.
- [8] K. Mase, "Recognition of facial expression from optical flow," *IEICE Transactions*, vol. E74, no. 10, pp. 3474–3483, 1991.
- [9] P. Ekman, "Facial expression and emotion," *American Psychologist*, vol. 8, no. 10, pp. 384–392, 1993.
- [10] J. F. Cohn, A. J. Zlochower, J. J. Lien, and T. Kanade, "Feature-point tracking by optical flow discriminates subtle differences in facial expression," *Automatic Face and Gesture Recognition*, pp. 396–401, 1998.
- [11] J. G. Daugman, "Complete discrete 2d gabor transform by neural networks for image analysis and compression," *Automatic IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 36, pp. 1169–1179, 1988.
- [12] M. Lyons and S. Akamatsu, "Coding facial expressions with gabor wavelets," in *Proc. of the Third IEEE International Conference on Automatic Face and Gesture Recognition*, (Nara, Japan), pp. 200–205, April 14 1998.
- [13] M. Lades, J. Vorbruggen, J. Buhmann, J. Lange, W. Konen, C. Malsburg, and R. Wurtz, "Distortion invariant object recognition in the dynamic link architecture," *IEEE Trans. On Computers*, vol. 42, pp. 300–311, March 1993.
- [14] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *Pattern Analysis and Machine Intelligence*, vol. 19, pp. 711–720, July 1997.

- [15] M. S. Bartlett and T. J. Sejnowski, "Viewpoint invariant face recognition using independent component analysis and attractor networks," *Advances in Neural Information Processing Systems*, vol. 9, pp. 817–823, 1997.
- [16] A. Lanitis, C. Taylor, and T. Cootes, "Automatic interpretation and coding of face images using flexible models," *Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 743–756, 1997.
- [17] C. Padgett and G. Cottrell, "Representing face images for emotion classification," in *Advances in Neural Information Processing Systems* (M. Mozer, M. Jordan, and T. Petsche, eds.), Cambridge, Mass: MIT Press, 1997.
- [18] N. Yrizarry, D. Matsumoto, and C. Wilson-Cohn, "American–japanese differences in multiscale intensity ratings of universal facial expressions of emotion," *Motivation and Emotion*, vol. 22, no. 4, pp. 315–327, 1998.
- [19] M. Pantic and A. Vinciarelli, "Implicit human centered tagging," *IEEE Signal Processing Magazine*, vol. 26, no. 6, 2009.
- [20] L. Sobel, "Camera models and machine perception," *PhD thesis*, 1970.
- [21] J. Prewitt, "Object enhancement and extraction," *Picture Process Psychopict*, pp. 75–149, 2009.
- [22] J. Y. Zhang, Y. Chen, and H. X. X, "Edge detection of images based on improved sobel operator and genetic algorithms," *Image Analysis and Signal Processing*, pp. 31–35, 2009.
- [23] D. C. M. and E. Hildreth, "Theory of edge detection," *Proceedings of The Royal Society*, vol. B275, pp. 199–210.

- [24] J. Canny, "A computational approach to edge detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 8, pp. 679–698, 1986.
- [25] A. Khashman and K. M. Curtis, "Neural networks arbitration for automatic edge detection of 3-dimensional objects," *Electronics, Circuits, and Systems*, vol. 1, pp. 49–52, 1996.
- [26] R. Deriche, "Fast algorithms for low-level vision," *IEEE Trans, PAMI*, vol. 12(A), pp. 78–87, 1990.
- [27] S. Lanser and W. Eckstein, "Eine modification des deriche-verfahrens zur kantendetektion," *DAGM Symposium*, pp. 151–158, 1991.
- [28] B. Jahne, H. Scharr, and S. Korgel, "Principles of filter design," in *Computer Vision and Applications* (B. Jahne, H. HauEecker, and P. GeiEJe, eds.), pp. 125–151, San Diego: Academic Press, 1999.
- [29] V. S. Nalwa and T. O. Binford, "On detecting edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, pp. 699–714, 1986.
- [30] P. J. Terry and D. Vu, "Edge detection using neural networks," *Signals, Systems and Computee*, vol. 1, pp. 391–395, 1993.
- [31] L. Zhai, S. Dong, and H. Ma, "Recent methods and applications on image edge detection," *Education Technology and Training. International Workshop on Geoscience and Remote Sensing*, vol. 1, pp. 332–335, 2008.
- [32] M. Sonka, V. Hlavac, and R. Boyle, "Image processing, analysis, and machine vision," pp. 90–92, Pacific Grove, CA: USA: Thomson Learning and PT Press, 1999.

- [33] T. Chen, Q. H. Wu, and R. Rahmani-Torkaman, "A pseudo top-hat mathematical morphological approach to edge detection in dark regions," *Pattern Recognition*, vol. 35, pp. 199–210, 2002.
- [34] N. E. Huang, Z. Shen, S. R. Long, M. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. of The Royal Society*, vol. 454, pp. 903–995, March 8 1998.
- [35] E. Delchelle, J. Lemoine, and O. Niang, "Empirical mode decomposition: An analytical approach for sifting process," *IEEE Signal Processing Letters*, vol. 12, pp. 764–767, November 2005.
- [36] Z. Lu, J. S. Smith, Q. H. Wu, and J. Fitch, "Empirical mode decomposition for power quality monitoring," *Transmission and Distribution Conference and Exhibition: Asia and Pacific*, vol. 1, pp. 1–5, 2005.
- [37] P. J. OONINCX, "Empirical mode decomposition: A new tool for s-wave detection," *Probability, networks and algorithms*, no. 3, pp. 1–14, 2002.
- [38] R. Liu, "Empirical mode decomposition: a useful technique for neuroscience," *Computational Journal Club, Univ. of California San Francisco*, January 2002.
- [39] N. E. Huang, M. L. Wu, W. Qu, S. R. Long, and S. P. Shen, "Applications of hilberthuang transform to non-stationary financial time series analysis," *Appl. Stochastic Models Bus. Ind.*, vol. 361, no. 19, pp. 245–268, 2003.
- [40] G. Rilling, P. Flandrin, and P. Goncalves, "Empirical mode decomposition as a filter bank," *IEEE Signal Process Letters*, vol. 11, no. 2, pp. 112–114, 2004.

- [41] P. Flandrin, P. Goncalves, and G. Rilling, “Detrending and denoising with emd,” *Proc. EUSIPCO*, pp. 1581–1584, 2004.
- [42] C. Damerval, S. Meignen, and C. Perrier, “A fast algorithm for bidimensional emd,” *IEEE Signal Processing Letters*, vol. 12, no. 10, pp. 701–704, 2005.
- [43] J. F. Sharif M.A.Bhuiyan, Reza R.Adhami, “Edge detection via a fast and adaptive bidimensional empirical mode decomposition,” *workshop on machine learning for signal processing (MLSP)*, pp. 199–204, 2008.
- [44] L. F. Chen and Y. S. Yen, “Taiwanese facial expression image database. retreat from <http://bml.ym.edu.tw/~download/html>,” *Brain Mapping Laboratory, Institute of Brain Science, National Yang-Ming University, Taipei, Taiwan*, 2007.

## **Appendix**

## APPENDIX A: COLLECTED DATA

Table A.1: Raw data of facial expression metrics tested on JAFFE

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
Model	Expressions	D1	D2	D3	D4
1	NE	53	21	14	37
	SA1	55		12	
	SA2	57		14	
	SA3	59		5	
	HA1	58		19	
	HA2	66		22	
	HA3	64		22	
	FE1	55		17	
	FE2	57		11	
	FE3	53		19	
	FE4	55		18	
	AN1	57	13		34
	AN2	56	13		29
	AN3	45	14		27
	DI1	52	20		33
	DI2	50	20		28
	DI3	52	20		31
	SU1	41	40		44
	SU2	59	27		43
	SU3	35	36		43
2	NE	51	24	17	38
	SA1	51		11	
	SA2	52		8	
	SA3	58		8	
	HA1	59		20	

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	HA2	61		21	
	HA3	64		24	
	FE1	49		19	
	FE2	47		16	
	FE3	49		19	
	FE4	56		19	
	DI1	50	26		36
	DI2	50	16		29
	DI3	58	13		29
	AN1	52	19		31
	AN2	50	19		24
	AN3	49	19		26
	SU1	56	24		49
	SU2	47	24		42
	SU3	43	32		43
<b>3</b>	<b>NE</b>	<b>47</b>	<b>17</b>	<b>10</b>	<b>36</b>
	SA1	52		6	
	SA2	50		5	
	SA3	49		8	
	HA1	69		19	
	HA2	72		18	
	HA3	70		16	
	FE1	50		16	
	FE2	48		14	
	FE3	49		12	
	DI1	48	17		31
	DI2	42	20		33
	DI3	47	19		32
	AN1	65	16		30
	AN2	61	13		30
	AN3	48	19		31
	SU1	42	31		40
	SU2	43	38		37
	SU3	42	34		37

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
4	NE	48	24	16	38
	SA1	46		8	
	SA2	41		7	
	SA3	45		7	
	HA1	64		25	
	HA2	66		23	
	HA3	65		20	
	FE1	53		17	
	FE2	52		16	
	FE3	53		12	
	DI1	45	16		23
	DI2	40	22		25
	AN1	39	22		28
	AN2	39	23		25
	AN3	47	26		27
	SU1	45	41		44
	SU2	41	40		46
	SU3	47	33		42
5	NE	52	19	14	33
	SA1	62		9	
	SA2	56		13	
	SA3	53		15	
	HA1	64		24	
	HA2	59		21	
	HA3	64		26	
	FE1	56		21	
	FE2	53		18	
	FE3	60		17	
	DI1	54	20		30
	DI2	54	19		30
	DI3	52	22		32
	AN1	49	21		25
	AN2	52	22		25
	AN3	48	19		26

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	SU1	35	48		46
	SU2	39	48		42
	SU3	47	36		42
<b>6</b>	<b>NE</b>	<b>53</b>	<b>20</b>	<b>18</b>	<b>34</b>
	SA1	55		9	
	SA2	54		6	
	SA3	55		13	
	HA1	67		27	
	HA2	63		22	
	HA3	62		25	
	FE1	53		13	
	FE2	53		13	
	FE3	55		14	
	DI1	51	19		34
	DI2	58	23		29
	DI3	51	24		33
	AN1	52	18		27
	AN2	50	17		31
	AN3	52	16		32
	SU1	60	28		43
	SU2	52	29		41
	SU3	48	23		39
<b>7</b>	<b>NE</b>	<b>49</b>	<b>28</b>	<b>20</b>	<b>36</b>
	SA1	50		12	
	SA2	50		13	
	SA3	51		19	
	HA1	74		21	
	HA2	73		22	
	HA3	58		13	
	FE1	50		19	
	FE2	52		17	
	FE3	50		14	
	DI1	52	24		40

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	DI2	48	21		40
	AN1	57	22		33
	AN2	54	21		35
	AN3	52	23		35
	SU1	46	36		41
	SU2	45	37		44
	SU3	47	33		46
<b>8</b>	<b>NE</b>	<b>51</b>	<b>21</b>	<b>15</b>	<b>32</b>
	SA1	52		12	
	SA2	52		13	
	SA3	55		13	
	HA1	56		17	
	HA2	57		22	
	HA3	57		21	
	FE1	50		18	
	FE2	48		15	
	FE3	46		21	
	DI1	48	21		32
	DI2	50	19		29
	DI3	49	24		31
	AN1	47	18		30
	AN2	46	17		31
	AN3	49	16		32
	SU1	40	32		36
	SU2	44	35		37
	SU3	50	33		36
<b>9</b>	<b>NE</b>	<b>49</b>	<b>21</b>	<b>15</b>	<b>34</b>
	SA1	54		8	
	SA2	52		7	
	SA3	51		10	
	HA1	53		16	
	HA2	55		16	
	HA3	57		20	

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	FE1	51		15	
	FE2	49		10	
	FE3	50		12	
	DI1	47	22		34
	DI2	48	18		37
	DI3	47	19		38
	AN1	49	19		32
	AN2	48	17		32
	AN3	52	16		31
	SU1	46	22		36
	SU2	48	22		38
	SU3	47	23		43

Table A.2: Raw data of facial expression metrics tested on TFEID

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
<b>Model</b>	<b>Expressions</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
f1	NE	74	37	20	46
	HA	89		36	
	SA	76		14	
	FE	79		16	
	DI	73	26		41
	AN	76	30		37
	SU	46	50		51
f2	NE	71	33	21	41
	HA	98		33	
	SA	73		20	
	FE	76		23	
	DI	60	25		40
	AN	73	24		39

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	SU	51	48		48
f3	NE	79	25	13	53
	HA	96		28	
	SA	80		6	
	FE	84		20	
	DI	61	19		28
	AN				
	SU	48	32		63
f4	NE	71	28	21	50
	HA	87		28	
	SA	76		16	
	FE	75		19	
	DI	60	23		37
	AN	74	20		32
	SU	48	45		56
f5	NE	64	27	16	43
	HA	83		32	
	SA	74		15	
	FE	74		17	
	DI	60	26		43
	AN	71	20		37
	SU	39	43		49
f6	NE	74	27	15	43
	HA	104		33	
	SA	79		13	
	FE	78		17	
	DI	76	27		41
	AN	68	24		36
	SU				
f8	NE	70	30	13	46
	HA	90		29	
	SA	82		11	
	FE	75		18	
	DI	71	17		41
	AN	66	19		39

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	SU	52	46		57
f9	NE	81	28	16	54
	HA	92		37	
	SA	78		12	
	FE	89		13	
	DI	64	20		42
	AN	75	16		42
	SU				
f10	NE	76	33	18	51
	HA	93		26	
	SA	86		6	
	FE	86		19	
	DI	63	22		38
	AN	75	14		31
	SU	46	48		55
f11	NE	77	38	21	44
	HA	99		29	
	SA	88		16	
	FE	91		26	
	DI	68	33		43
	AN	71	20		45
	SU	51	44		45
f12	NE	80	34	18	40
	HA	102		31	
	SA	81		8	
	FE	81		23	
	DI	77	31		51
	AN	76	23		37
	SU	55	48		51
f13	NE	74	28	19	47
	HA	88		32	
	SA	80		13	
	FE	82		16	
	DI	70	32		42
	AN	72	20		42

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	SU	54	53		59
f14	NE	74	32	19	43
	HA	107		40	
	SA	75		16	
	FE	74		21	
	DI	69	33		35
	AN	83	27		39
	SU	51	60		53
f15	NE	77	32	17	48
	HA	100		32	
	SA	78		15	
	FE	91		14	
	DI	73	30		31
	AN	83	24		39
	SU				
f16	NE	81	26	18	48
	HA	92		38	
	SA	79		13	
	FE	75		10	
	DI	69	33		39
	AN	83	25		37
	SU	58	39		60
f17	NE	68	35	27	46
	HA	93		31	
	SA	78		13	
	FE	78		22	
	DI	76	28		46
	AN				
	SU	54	56		58
f18	NE	71	32	23	47
	HA	95		33	
	SA	75		12	
	FE	74		20	
	DI	61	29		34
	AN	74	22		39

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	SU	52	50		61
f19	NE	69	34	26	42
	HA	92		37	
	SA	80		16	
	FE	79		22	
	DI	73	36		32
	AN	79	20		34
	SU	54	50		53
f20	NE	72	26	21	41
	HA	95		31	
	SA	87		13	
	FE	90		19	
	DI	80	31		41
	AN	77	26		40
	SU	44	53		56
f21	NE	72	27	15	49
	HA	83		35	
	SA	78		12	
	FE	79		25	
	DI	63	32		43
	AN	63	23		41
	SU	43	47		57
m1	NE	81	25	13	38
	HA	92		23	
	SA	82		14	
	FE	87		19	
	DI	78	23		31
	AN				
	SU	64	38		44
m2	NE	79	23	12	38
	HA	98		26	
	SA				
	FE	80		13	
	DI	70	33		29
	AN				

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	SU	50	56		44
<b>m3</b>	<b>NE</b>	<b>78</b>	<b>20</b>	<b>15</b>	<b>42</b>
	HA	84		31	
	SA	87		9	
	FE	83		15	
	DI	83	25		34
	AN				
	SU	66	32		47
<b>m4</b>	<b>NE</b>	<b>78</b>	<b>30</b>	<b>15</b>	<b>42</b>
	HA	93		28	
	SA	78		7	
	FE	89		13	
	DI	77	39		41
	AN	76	18		34
	SU				
<b>m5</b>	<b>NE</b>	<b>66</b>	<b>20</b>	<b>13</b>	<b>35</b>
	HA	89		18	
	SA	80		10	
	FE	80		16	
	DI	65	26		31
	AN	66	12		29
	SU	49	42		40
<b>m6</b>	<b>NE</b>	<b>73</b>	<b>28</b>	<b>11</b>	<b>39</b>
	HA	94		24	
	SA	73		10	
	FE	70		20	
	DI	72	29		35
	AN				
	SU	52	54		51
<b>m7</b>	<b>NE</b>	<b>73</b>	<b>29</b>	<b>18</b>	<b>39</b>
	HA	98		23	
	SA	82		9	
	FE	81		24	
	DI	63	26		35
	AN	79	13		38

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	SU	46	36		58
<b>m8</b>	<b>NE</b>	<b>81</b>	<b>22</b>	<b>14</b>	<b>38</b>
	HA	103		25	
	SA	83		7	
	FE	66		19	
	DI	70	38		39
	AN	73	24		32
	SU	52	42		48
<b>m9</b>	<b>NE</b>	<b>71</b>	<b>20</b>	<b>14</b>	<b>39</b>
	HA	92		21	
	SA	63		6	
	FE	71		12	
	DI	69	27		26
	AN	70	12		27
	SU	47	52		53
<b>m10</b>	<b>NE</b>	<b>73</b>	<b>22</b>	<b>12</b>	<b>32</b>
	HA	102		32	
	SA	78		10	
	FE	81		15	
	DI	79	20		30
	AN	72	16		38
	SU	52	48		49
<b>m11</b>	<b>NE</b>	<b>66</b>	<b>26</b>	<b>16</b>	<b>46</b>
	HA	78		10	
	SA	76		13	
	FE	78		15	
	DI	69	31		27
	AN	65	17		36
	SU	56	52		57
<b>m12</b>	<b>NE</b>	<b>78</b>	<b>25</b>	<b>12</b>	<b>48</b>
	HA	88		37	
	SA	88		6	
	FE	86		14	
	DI	72	28		41
	AN	74	22		32

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	SU	58	55		61
<b>m13</b>	<b>NE</b>	<b>77</b>	<b>29</b>	<b>13</b>	<b>44</b>
	HA	103		32	
	SA	84		2	
	FE	87		15	
	DI	65	33		25
	AN	77	11		34
	SU	53	52		55
<b>m14</b>	<b>NE</b>	<b>84</b>	<b>31</b>	<b>20</b>	<b>44</b>
	HA	95		29	
	SA	89		3	
	FE	91		19	
	DI	70	30		41
	AN	76	15		26
	SU	60	43		57
<b>m15</b>	<b>NE</b>	<b>70</b>	<b>20</b>	<b>10</b>	<b>42</b>
	HA	90		30	
	SA	76		8	
	FE	77		20	
	DI	61	21		35
	AN	72	17		40
	SU	46	52		60
<b>m16</b>	<b>NE</b>	<b>66</b>	<b>23</b>	<b>20</b>	<b>38</b>
	HA	93		30	
	SA	78		6	
	FE	79		12	
	DI	54	24		28
	AN	72	21		28
	SU	46	58		50
<b>m17</b>	<b>NE</b>	<b>74</b>	<b>21</b>	<b>14</b>	<b>30</b>
	HA	91		30	
	SA	80		11	
	FE	91		13	
	DI	65	32		38
	AN	68	17		28

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	SU	45	56		53
<b>m18</b>	<b>NE</b>	<b>75</b>	<b>25</b>	<b>15</b>	<b>38</b>
	HA	89		18	
	SA	67		13	
	FE	70		15	
	DI	59	22		28
	AN	59	12		29
	SU	51	49		56
<b>m19</b>	<b>NE</b>	<b>73</b>	<b>25</b>	<b>13</b>	<b>33</b>
	HA	91		27	
	SA	90		4	
	FE	73		18	
	DI	67	26		40
	AN	79	21		24
	SU	44	50		51
<b>m20</b>	<b>NE</b>	<b>81</b>	<b>19</b>	<b>14</b>	<b>34</b>
	HA	103		34	
	SA	82		13	
	FE	82		12	
	DI	75	32		47
	AN	91	18		33
	SU	54	53		36

Table A.3: Raw data of facial expression metrics blind detection tested on JAFFE

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
Model	Expressions	D1	D2	D3	D4
1	NE	53	21	14	37
	SA1	51		13	
	SA2	56		12	

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	SA3	59		5	
	HA1	58		16	
	HA2	65		21	
	HA3	63		20	
	FE1	53		15	
	FE2	54		11	
	FE3	50		15	
	FE4	55		14	
	AN1	49	12		33
	AN2	55	22		26
	AN3	48	19		27
	DI1	52	21		33
	DI2	50	22		25
	DI3	52	20		27
	SU1	35	24		40
	SU2	59	27		42
	SU3	40	24		41
<b>2</b>	<b>NE</b>	<b>51</b>	<b>24</b>	<b>17</b>	<b>38</b>
	SA1	51		10	
	SA2	53		14	
	SA3	58		14	
	HA1	59		19	
	HA2	61		22	
	HA3	63		17	
	FE1	49		13	
	FE2	47		13	
	FE3	49		14	
	FE4	51		16	
	DI1	52	27		61
	DI2	55	30		58
	DI3	59	30		55
	AN1	52	29		30
	AN2	51	30		28
	AN3	50	28		34
	SU1	57	15		48

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	SU2	47	22		42
	SU3	43	21		46
3	NE	47	17	10	36
	SA1	51		8	
	SA2	49		8	
	SA3	47		8	
	HA1	63		16	
	HA2	64		18	
	HA3	66		13	
	FE1	49		13	
	FE2	48		14	
	FE3	48		15	
	DI1	44	33		38
	DI2	40	20		35
	DI3	47	18		32
	AN1	64	11		33
	AN2	62	24		27
	AN3	47	23		33
	SU1	42	33		42
	SU2	43	33		37
	SU3	42	33		35
4	NE	48	24	16	38
	SA1	45		8	
	SA2	40		8	
	SA3	44		9	
	HA1	51		27	
	HA2	51		28	
	HA3	51		25	
	FE1	51		18	
	FE2	51		20	
	FE3	50		16	
	DI1	44	32		50
	DI2	40	28		28
	AN1	37	28		31
	AN2	38	22		28

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	AN3	47	22		47
	SU1	41	28		41
	SU2	41	28		41
	SU3	40	29		41
<b>5</b>	<b>NE</b>	<b>47</b>	<b>19</b>	<b>14</b>	<b>33</b>
	SA1	52		16	
	SA2	52		16	
	SA3	50		14	
	HA1	52		22	
	HA2	52		23	
	HA3	52		25	
	FE1	52		12	
	FE2	46		16	
	FE3	46		17	
	DI1	50	19		33
	DI2	52	22		33
	DI3	50	21		33
	AN1	52	28		28
	AN2	50	33		33
	AN3	50	17		27
	SU1	27	42		37
	SU2	36	29		42
	SU3	43	36		42
<b>6</b>	<b>NE</b>	<b>49</b>	<b>20</b>	<b>18</b>	<b>34</b>
	SA1	50		6	
	SA2	49		14	
	SA3	51		12	
	HA1	52		22	
	HA2	52		24	
	HA3	52		24	
	FE1	50		3	
	FE2	50		9	
	FE3	52		4	
	DI1	45	14		29
	DI2	52	15		37

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	DI3	48	21		31
	AN1	49	18		28
	AN2	48	17		31
	AN3	50	16		29
	SU1	52	22		43
	SU2	52	19		37
	SU3	51	12		39
<b>7</b>	<b>NE</b>	<b>49</b>	<b>28</b>	<b>16</b>	<b>36</b>
	SA1	46		12	
	SA2	48		14	
	SA3	51		12	
	HA1	54		14	
	HA2	54		16	
	HA3	55		16	
	FE1	49		14	
	FE2	51		16	
	FE3	49		18	
	DI1	50	38		42
	DI2	52	20		42
	AN1	54	21		41
	AN2	52	18		40
	AN3	51	23		39
	SU1	46	30		45
	SU2	44	30		43
	SU3	45	31		46
<b>8</b>	<b>NE</b>	<b>51</b>	<b>25</b>	<b>19</b>	<b>32</b>
	SA1	51		17	
	SA2	49		16	
	SA3	47		18	
	HA1	54		26	
	HA2	54		21	
	HA3	54		14	
	FE1	54		22	
	FE2	48		19	
	FE3	46		22	

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	DI1	49	31		31
	DI2	51	34		36
	DI3	52	31		33
	AN1	43	27		31
	AN2	54	29		35
	AN3	43	33		29
	SU1	40	27		36
	SU2	43	35		37
	SU3	45	32		36
9	NE	49	21	15	34
	SA1	52		9	
	SA2	52		7	
	SA3	51		40	
	HA1	53		14	
	HA2	54		14	
	HA3	54		18	
	FE1	51		16	
	FE2	47		14	
	FE3	46		15	
	DI1	47	18		32
	DI2	47	21		34
	DI3	52	26		37
	AN1	52	38		32
	AN2	48	17		31
	AN3	50	18		34
	SU1	47	21		36
	SU2	46	33		38
	SU3	47	29		40

Table A.4: Raw data of facial expression metrics blind detection tested on TFEID

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
f1	NE	74	33	20	50
	HA	80		31	
	SA	75		12	
	FE	78		14	
	DI	86	35		33
	AN	76	25		35
	SU	45	43		35
f2	NE	71	37	21	43
	HA	77		26	
	SA	72		19	
	FE	78		22	
	DI	60	37		35
	AN	72	25		37
	SU	66	43		46
f3	NE	85	41	13	44
	HA	87		21	
	SA	85		11	
	FE	83		9	
	DI	76	39		44
	AN				
	SU	50	41		60
f4	NE	72	27	21	46
	HA	83		30	
	SA	78		13	
	FE	73		19	
	DI	73	35		38
	AN	77	29		40
	SU	48	38		55
f5	NE	65	34	16	47
	HA	83		30	
	SA	74		16	
	FE	75		14	
	DI	55	39		54

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	AN	84	25		37
	SU	44	39		42
<b>f6</b>	<b>NE</b>	<b>74</b>	<b>35</b>	<b>15</b>	<b>37</b>
	HA	85		22	
	SA	75		16	
	FE	85		25	
	DI	83	39		49
	AN	85	39		31
	SU				
<b>f8</b>	<b>NE</b>	<b>73</b>	<b>37</b>	<b>16</b>	<b>52</b>
	HA	85		29	
	SA	81		14	
	FE	73		14	
	DI	85	24		43
	AN	85	31		38
	SU	52	37		57
<b>f9</b>	<b>NE</b>	<b>79</b>	<b>25</b>	<b>16</b>	<b>55</b>
	HA	85		35	
	SA	79		5	
	FE	85		8	
	DI	82	39		31
	AN	74	18		41
	SU				
<b>f10</b>	<b>NE</b>	<b>82</b>	<b>21</b>	<b>18</b>	<b>49</b>
	HA	85		35	
	SA	85		4	
	FE	85		11	
	DI	70	28		30
	AN	78	21		42
	SU	49	39		56
<b>f11</b>	<b>NE</b>	<b>76</b>	<b>35</b>	<b>18</b>	<b>51</b>
	HA	85		28	
	SA	84		25	
	FE	85		20	
	DI	72	36		47
	AN	75	20		47
	SU	55	39		65
<b>f12</b>	<b>NE</b>	<b>83</b>	<b>35</b>	<b>18</b>	<b>39</b>

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	HA	89		35	
	SA	85		21	
	FE	81		37	
	DI	70	39		40
	AN	79	34		60
	SU	57	39		57
<b>f13</b>	<b>NE</b>	<b>76</b>	<b>32</b>	<b>18</b>	<b>42</b>
	HA	85		32	
	SA	83		20	
	FE	85		20	
	DI	68	39		31
	AN	77	19		26
	SU	53	37		55
<b>f14</b>	<b>NE</b>	<b>74</b>	<b>37</b>	<b>23</b>	<b>38</b>
	HA	85		37	
	SA	72		15	
	FE	74		21	
	DI	67	39		22
	AN	85	30		38
	SU	53	39		50
<b>f15</b>	<b>NE</b>	<b>77</b>	<b>39</b>	<b>18</b>	<b>47</b>
	HA	85		34	
	SA	80		14	
	FE	85		28	
	DI	77	30		40
	AN	85	17		29
	SU				
<b>f16</b>	<b>NE</b>	<b>84</b>	<b>37</b>	<b>18</b>	<b>49</b>
	HA	85		35	
	SA	85		16	
	FE	84		12	
	DI	64	39		36
	AN	85	39		46
	SU	52	39		60
<b>f17</b>	<b>NE</b>	<b>74</b>	<b>25</b>	<b>27</b>	<b>40</b>
	HA	85		39	
	SA	82		22	
	FE	80		18	

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	DI	64	21		37
	AN				
	SU	50	39		58
f18	NE	76	39	16	51
	HA	85		31	
	SA	76		12	
	FE	80		18	
	DI	61	38		26
	AN	72	20		40
	SU	46	39		57
f19	NE	72	39	21	41
	HA	85		25	
	SA	85		19	
	FE	81		17	
	DI	85	39		41
	AN	80	24		35
	SU	50	39		56
f20	NE	83	35	21	39
	HA	85		39	
	SA	82		10	
	FE	82		7	
	DI	84	35		47
	AN	85	35		41
	SU	44	39		64
f21	NE	73	27	15	51
	HA	85		34	
	SA	71		26	
	FE	81		25	
	DI	63	30		39
	AN	65	29		38
	SU	42	37		51
m1	NE	83	30	13	37
	HA	85		23	
	SA	82		25	
	FE	85		23	
	DI	78	36		29
	AN				
	SU	64	37		61

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
m2	NE	80	35	12	38
	HA	85		17	
	SA				
	FE	83		30	
	DI	71	34		39
	AN				
	SU	70	39		66
m3	NE	79	37	20	61
	HA	84		31	
	SA	85		8	
	FE	83		21	
	DI	84	27		62
	AN				
	SU	68	36		68
m4	NE	80	38	19	43
	HA	85		34	
	SA	77		23	
	FE	85		23	
	DI	85	38		34
	AN	81	38		52
	SU				
m5	NE	66	17	13	39
	HA	85		34	
	SA	80		13	
	FE	80		18	
	DI	67	27		39
	AN	72	13		58
	SU	50	37		52
m6	NE	74	29	15	56
	HA	85		27	
	SA	73		10	
	FE	71		19	
	DI	75	26		41
	AN				
	SU	61	38		53
m7	NE	73	29	18	39
	HA	85		24	
	SA	81		24	

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	FE	84		24	
	DI	63	25		29
	AN	85	31		50
	SU	55	34		65
<b>m8</b>	<b>NE</b>	<b>83</b>	<b>35</b>	<b>14</b>	<b>34</b>
	HA	85		24	
	SA	83		25	
	FE	66		11	
	DI	71	38		50
	AN	83	31		39
	SU	53	38		62
<b>m9</b>	<b>NE</b>	<b>73</b>	<b>29</b>	<b>14</b>	<b>39</b>
	HA	85		25	
	SA	66		18	
	FE	76		15	
	DI	71	25		35
	AN	75	20		35
	SU	48	34		62
<b>m10</b>	<b>NE</b>	<b>73</b>	<b>31</b>	<b>12</b>	<b>32</b>
	HA	85		34	
	SA	78		20	
	FE	85		20	
	DI	81	31		39
	AN	76	26		36
	SU	59	37		70
<b>m11</b>	<b>NE</b>	<b>80</b>	<b>36</b>	<b>13</b>	<b>55</b>
	HA	79		13	
	SA	83		16	
	FE	84		16	
	DI	79	34		28
	AN	71	36		38
	SU	61	34		67
<b>m12</b>	<b>NE</b>	<b>79</b>	<b>29</b>	<b>13</b>	<b>46</b>
	HA	85		18	
	SA	85		21	
	FE	82		20	
	DI	85	39		50
	AN	85	39		37

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	SU	60	39		63
<b>m13</b>	NE	80	32	12	45
	HA	85		24	
	SA	85		4	
	FE	85		19	
	DI	66	35		23
	AN	81	12		33
	SU	63	39		55
<b>m14</b>	NE	83	39	20	43
	HA	85		29	
	SA	85		4	
	FE	85		23	
	DI	74	39		42
	AN	82	32		40
	SU	68	39		64
<b>m15</b>	NE	72	31	12	48
	HA	85		32	
	SA	78		8	
	FE	78		19	
	DI	76	26	39	
	AN	85	30	44	
	SU	49	39	66	
<b>m16</b>	NE	71	39	20	36
	HA	85		39	
	SA	78		9	
	FE	82		10	
	DI	56	36		33
	AN	76	28		27
	SU	47	39		51
<b>m17</b>	NE	74	22	14	34
	HA	85		18	
	SA	84		15	
	FE	85		20	
	DI	66	29		40
	AN	85	27		37
	SU	56	38		68
<b>m18</b>	NE	75	25	15	38
	HA	85		22	

*Continued on next page*

<b>Model</b>	<b>Expression</b>	<b>D1</b>	<b>D2</b>	<b>D3</b>	<b>D4</b>
	SA	84		15	
	FE	78		21	
	DI	61	24		25
	AN	67	38		40
	SU	55	34		70
<b>m19</b>	<b>NE</b>	<b>75</b>	<b>27</b>	<b>13</b>	<b>32</b>
	HA	85		15	
	SA	85		11	
	FE	72		15	
	DI	68	34		43
	AN	85	28		37
	SU	50	38		65
<b>m20</b>	<b>NE</b>	<b>81</b>	<b>19</b>	<b>14</b>	<b>50</b>
	HA	85		23	
	SA	84		14	
	FE	85		23	
	DI	78	26		47
	AN	85	25		55
	SU	61	38		70